

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:17:38 ; Search time 9447.36 Seconds  
(without alignments)  
11696.741 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 cttgaagacactggatttc.....cctttgtgaaaaaaaaa 3797

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3388.8	89.2	6028	9	AF309033	AF309033 Homo sapi
2	3388.8	89.2	6123	6	AX057579	AX057579 Sequence
3	3387.2	89.2	3815	9	AF329696	AF329696 Homo sapi
4	3387.2	89.2	4127	6	AX062241	AX062241 Sequence
5	3387.2	89.2	4406	6	AX062247	AX062247 Sequence
6	3387.2	89.2	4992	6	AX062275	AX062275 Sequence
7	3387.2	89.2	5002	6	AX062273	AX062273 Sequence
8	3387.2	89.2	5810	9	AF264912	AF264912 Homo sapi
9	3387.2	89.2	6189	9	AF438201	AF438201 Homo sapi
10	3386	89.2	3394	6	AX133694	AX133694 Sequence
11	3382.8	89.1	4512	6	AX076213	AX076213 Sequence
12	3338.4	87.9	4295	9	AF305081	AF305081 Homo sapi
13	3300.4	86.9	3501	9	AF342982	AF342982 Homo sapi
14	3297	86.8	3400	6	AX029397	AX029397 Sequence
15	3296.2	86.8	3508	6	AX062142	AX062142 Sequence
16	3291.8	86.7	3498	6	AX076214	AX076214 Sequence
17	2774.6	73.1	5005	6	AX052530	AX052530 Sequence
18	2439.8	64.3	2971	6	AX062233	AX062233 Sequence
19	2439.8	64.3	3353	6	AX062229	AX062229 Sequence
20	2357.8	62.1	3799	6	AX062231	AX062231 Sequence
21	1922.4	50.6	4130	9	HS804989	AL833676 Homo sapi
22	1813.4	47.8	3984	6	AX062144	AX062144 Sequence
23	1813.4	47.8	4134	6	AX467061	AX467061 Sequence
24	1813.4	47.8	4134	9	AF082556	AF082556 Homo sapi
25	1794.8	47.3	3254	6	AX057580	AX057580 Sequence
26	1794.8	47.3	8901	9	AF082557	AF082557 Homo sapi
27	1449.6	38.2	4491	6	AX467067	AX467067 Sequence
28	1359.4	35.8	1568	6	AX062239	AX062239 Sequence
29	1297.8	34.2	4657	6	AX467069	AX467069 Sequence
30	1172.2	30.9	1242	6	AX052522	AX052522 Sequence
31	1168.4	30.8	3308	6	AX062318	AX062318 Sequence
32	1084.8	28.6	1103	6	AX062300	AX062300 Sequence
33	1076.8	28.4	3006	9	AX023746	AX023746 Homo sapi
34	1008.8	26.6	5482	3	AF132196	AF132196 Drosophil
35	1008.8	26.6	5482	6	AX062279	AX062279 Sequence
36	994	26.2	2409	6	AX429577	AX429577 Sequence
37	990	26.1	1007	6	AX052525	AX052525 Sequence
38	748.6	19.7	48169	2	AC017396	AC017396 Drosophil
39	748.6	19.7	167396	3	AC010579	AC010579 Drosophil
40	748.6	19.7	225478	3	AE003752	AE003752 Drosophil
41	574	15.1	582	6	AX062186	AX062186 Sequence
42	459	12.1	460	6	AX062146	AX062146 Sequence
43	438.8	11.6	142085	9	AL359707	AL359707 Human DNA
44	396.8	10.5	402	6	AX201878	AX201878 Sequence
45	359.6	9.5	1691	6	AX062182	AX062182 Sequence

# ALIGNMENTS

RESULT 1  
AF309033  
LOCUS AF309033 6028 bp mRNA linear PRI 31-JUL-2001  
DEFINITION Homo sapiens tankyrase-2 (TNKS-2) mRNA, complete cds.  
ACCESSION AF309033  
VERSION AF309033.1 GI:15042551  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 6028)  
AUTHORS Chi,N.-W. and Lodish,H.F.  
TITLE A novel homolog of Tankyrase  
JOURNAL Unpublished

Pred. No. is the number of results predicted by chance to have a





Db	1829	CTCTACACCACTTCAATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT	1888
Qy	1843	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGCTTGTACTCTTGCACAATGC	1902
Db	1889	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGAGCCCTTGTACCTTTTGACAATGC	1948
Qy	1903	ATGTTCTTATGGACATATGAAGTTGCGAGAAGCTTCTGTTTAAACATGGACAGTAGTTAA	1962
Db	1949	ATGTTCTTATGGACATATGAAGTTGCGAGAAGCTTCTGTTTAAACATGGACAGTAGTTAA	2008
Qy	1963	TGTAGCTGATTTATGGAATTTTACACCTTTACATGAACGACGACCAAAAGAAAATATGA	2022
Db	2009	TGTAGCTGATTTATGGAATTTTACACCTTTACATGAACGACGACCAAAAGAAAATATGA	2068
Qy	2023	AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTACCAAAAAAACAGGATGGAAA	2082
Db	2069	AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTACCAAAAAAACAGGATGGAAA	2128
Qy	2083	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGA	2142
Db	2129	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGA	2188
Qy	2143	TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTC	2202
Db	2189	TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTC	2248
Qy	2203	TCCTGATAATGTAAATTCGCCGATACCAAGGCAGACATTCACACCTTTACATTTAGC	2262
Db	2249	TCCTGATAATGTAAATTCGCCGATACCAAGGCAGACATTCACACCTTTACATTTAGC	2308
Qy	2263	AGCTGGTTATATTAATTTAGAAGTTGCGAGATTTTGTTCACACGAGGAGCTGATGTAA	2322
Db	2309	AGCTGGTTATATTAATTTAGAAGTTGCGAGATTTTGTTCACACGAGGAGCTGATGTAA	2368
Qy	2323	TGCCCAAGACAAGGAGGACTTATTCCTTTACATAATGCAGCATCTACGGGCATGTAGA	2382
Db	2369	TGCCCAAGACAAGGAGGACTTATTCCTTTACATAATGCAGCATCTACGGGCATGTAGA	2428
Qy	2383	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTT	2442
Db	2429	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTT	2488
Qy	2443	CACACCTTTGCAGAGCAGAGCCCAAAAGGAGCAGAACACAGCTTTTGTGTTGCTAGC	2502
Db	2489	CACACCTTTGCAGAGCAGAGCCCAAAAGGAGCAGAACACAGCTTTTGTGTTGCTAGC	2548
Qy	2503	CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAACACCTTTAGATTTAGTTTC	2562
Db	2549	CCATGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAACACCTTTAGATTTAGTTTC	2608
Qy	2563	ACGGGATGATGACGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCCCTCTTG	2622
Db	2609	ACGGGATGATGACGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCCCTCTTG	2668
Qy	2623	TTTACAAGCCTCAAGTGTCAATGTGTGAGAGCCAGGACCACTGCGATGCTCTCTC	2682
Db	2669	TTTACAAGCCTCAAGTGTCAATGTGTGAGAGCCAGGACCACTGCGATGCTCTCTC	2728
Qy	2683	TTTCAAGTCCATCTAGCCCATCAAGCCCTTCTGACGACGAGCTTGAACAACCTTATCTGG	2742
Db	2729	TTTCAAGTCCATCTAGCCCATCAAGCCCTTCTGACGACGAGCTTGAACAACCTTATCTGG	2788
Qy	2743	GAGTTTTTCAGAACTGTCTTCACTAGTATTCTCAAGTGAACAGAGGGTGTCTCCAGTTT	2802
Db	2789	GAGTTTTTCAGAACTGTCTTCACTAGTATTCTCAAGTGAACAGAGGGTGTCTCCAGTTT	2848
Qy	2803	GGAGAAAAGAGGTTCCAGAGTAGATTTTACATAACTTCAATTCGTAAAGGAATCTTGG	2862
Db	2849	GGAGAAAAGAGGTTCCAGAGTAGATTTTACATAACTTCAATTCGTAAAGGAATCTTGG	2908
Qy	2863	ACTTCAGCACCTTAATGGATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA	2922
Db	2909	ACTTCAGCACCTTAATGGATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA	2968

REFERENCE 1 (bases 1 to 6123)  
AUTHORS Chi.N.W. and Lodish.H.F.  
TITLE A novel insulin signaling molecule  
JOURNAL Patent: WO 0077225-A 2 21-DEC-2000;  
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; THE GENERAL  
HOSPITAL CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..6123  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1758 a 1189 c 1358 g 1818 t  
ORIGIN  
Query Match 89.2%; Score 3388.8; DB 6; Length 6123;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 403 AGGTTTGGCGGAAGAGCTAGTTCGAATATTTGCTTCAGATGGTGCAGGTGTCGAAGC 462  
DB 449 AGGTTTGGCGGAAGAGAGCTAGTTCGAATATTTGCTTCAGATGGTGCAGGTGTCGAAGC 508  
QY 463 ACCTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 522  
DB 509 ACCTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 568  
QY 523 AGTCAATCTCCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 582  
DB 569 AGTCAATCTCCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC 628  
QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATTTGGCTTGTGCTGTACAGCA 642  
DB 629 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATTTGGCTTGTGCTGTACAGCA 688  
QY 643 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTACAGATCC 702  
DB 689 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTACAGATCC 748  
QY 703 ATCTGCAAGAGAGCTTACTGCTGAATATAGAAAGATGAAGTCTAGAAAGTGCCAG 762  
DB 749 ATCTGCAAGAGAGCTTACTGCTGAATATAGAAAGATGAAGTCTAGAAAGTGCCAG 808  
QY 763 GAGTGGCAATCAAGAAATATGATGGCTTACTCACACCATTTAAATGTCACATGCCACGC 822  
DB 809 GAGTGGCAATCAAGAAATATGATGGCTTACTCACACCATTTAAATGTCACATGCCACGC 868  
QY 823 AAGTGATGGCAGAAAGTCACTCCATTTACATTTGGCAGCAGGATATACAGATTAAGAT 882  
DB 869 AAGTGATGGCAGAAAGTCACTCCATTTACATTTGGCAGCAGGATATACAGATTAAGAT 928  
QY 883 TGTACAGCTGTTACTGCAACATGGAGTGTATGCTCATGCTAAAGATTAAGGTGATCTGGT 942  
DB 929 TGTACAGCTGTTACTGCAACATGGAGTGTATGCTCATGCTAAAGATTAAGGTGATCTGGT 988  
QY 943 ACCATTACAAATGCCCTGTTCTTATAGTCAATTAAGATTAAGTCAAGTCTTTGGTCAAGCA 1002  
DB 989 ACCATTACAAATGCCCTGTTCTTATAGTCAATTAAGATTAAGTCAAGTCTTTGGTCAAGCA 1048  
QY 1003 TGGTGCCTGTGTAATGCAATGGACTTTGGCAATTTCACTCTCTTCATGAGGCAGCTTC 1062  
DB 1049 TGGTGCCTGTGTAATGCAATGGACTTTGGCAATTTCACTCTCTTCATGAGGCAGCTTC 1108  
QY 1063 TAAGAACAGGTTTGAAGTATGTTCTTCTTCTTAAAGTATGGTGCAGACCCCAACACTGCT 1122  
DB 1109 TAAGAACAGGTTTGAAGTATGTTCTTCTTCTTAAAGTATGGTGCAGACCCCAACACTGCT 1168  
QY 1123 CAATTGTCACAATTAAGTGTCTATAGACTTTGGCTCCCAACACCAAGTTTAAAGAAAGATT 1182  
DB 1169 CAATTGTCACAATTAAGTGTCTATAGACTTTGGCTCCCAACACCAAGTTTAAAGAAAGATT 1228  
QY 1183 AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG 1242  
DB 1229 AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG 1288

QY 1243 AATCAAAACACATCTCTCTGGAATGCTGAATTTCAAGCATCTTCAACACATGAAC 1302  
DB 1289 AATCAAAACACATCTCTCTGGAATGCTGAATTTCAAGCATCTTCAACACATGAAC 1348  
QY 1303 AGCATTTGCTTGTGCTGCTCCTCATATCCCAAAAGAAAGCAAAATATGTGAATGTT 1362  
DB 1349 AGCATTTGCTTGTGCTGCTCCTCATATCCCAAAAGAAAGCAAAATATGTGAATGTT 1408  
QY 1363 GCTAAGAAAGAGCAAAACATCAATGAAAGACTTAAGAAATCTTGACTCTCTGCACGT 1422  
DB 1409 GCTAAGAAAGAGCAAAACATCAATGAAAGACTTAAGAAATCTTGACTCTCTGCACGT 1468  
QY 1423 GGCATCTGAGAAAGCTCATTAATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482  
DB 1469 GGCATCTGAGAAAGCTCATTAATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT 1528  
QY 1483 TAATGCTCTGGATAAATCTTGGTTCAGACTTCTTACACAGAGCTGCATATTTGGTCTATCT 1542  
DB 1529 TAATGCTCTGGATAAATCTTGGTTCAGACTTCTTACACAGAGCTGCATATTTGGTCTATCT 1588  
QY 1543 ACAAACTGCCGCTACTCTGAGCTATGGTGTGATCTTAACATATATCCCTTCAGGG 1602  
DB 1589 ACAAACTGCCGCTACTCTGAGCTATGGTGTGATCTTAACATATATCCCTTCAGGG 1648  
QY 1603 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCACAGAGGTATCTC 1662  
DB 1649 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCACAGAGGTATCTC 1708  
QY 1663 ATTAGGTAATTCAGAGCAGACAGACAAATTTGCTTGAAGCTGCAAGAGCTGGAGATCTCGA 1722  
DB 1709 ATTAGGTAATTCAGAGCAGACAGACAAATTTGCTTGAAGCTGCAAGAGCTGGAGATCTCGA 1768  
QY 1723 AACTGTAAAAAATCTGTACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGCTCA 1782  
DB 1769 AACTGTAAAAAATCTGTACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGCTCA 1828  
QY 1783 GTCTACACCCTTCAATTTTCAGCTGGGTATACAGAGTGTCCGTGTGGAATATCTGCT 1842  
DB 1829 GTCTACACCCTTCAATTTTCAGCTGGGTATACAGAGTGTCCGTGTGGAATATCTGCT 1888  
QY 1843 ACAGCATGGAGCTGATGCTGATGCTAAAGATAAGAGAGGCTTGTACCTTTGCAACAATGC 1902  
DB 1889 ACAGCATGGAGCTGATGCTGATGCTAAAGATAAGAGAGGCTTGTACCTTTGCAACAATGC 1948  
QY 1903 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGAGCAGTAGTTAA 1962  
DB 1949 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTTCTTTTAAACATGAGCAGTAGTTAA 2008  
QY 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAAGCAGCAGCAAAAGAAATATGA 2022  
DB 2009 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAAGCAGCAGCAAAAGAAATATGA 2068  
QY 2023 AATTGTCAAACTTCTGCTCCAGCATGCTGAGACCCCTACCAAAAAACACAGGATGAAA 2082  
DB 2069 AATTGTCAAACTTCTGCTCCAGCATGCTGAGACCCCTACCAAAAAACACAGGATGAAA 2128  
QY 2083 TACTCTCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA 2142  
DB 2129 TACTCTCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA 2188  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGGAAGATTTGCTCTTC 2202  
DB 2189 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGGAAGATTTGCTCTTC 2248  
QY 2203 TCCTGTAAATGTAATTTGCCGCTATACCAAGGAGAGATTCACACCTTTTACATTTAGC 2262  
DB 2249 TCCTGTAAATGTAATTTGCCGCTATACCAAGGAGAGATTCACACCTTTTACATTTAGC 2308  
QY 2263 AGCTGGTTATTAATTTTAGAAGTTGAGATATTTGTACACACGAGCTGATGTTGAA 2322  
DB 2309 AGCTGGTTATTAATTTTAGAAGTTGAGATATTTGTACACACGAGCTGATGTTGAA 2368  
QY 2323 TGCCCAAGACAAAGAGGAGGACTTTATTCCTTTTACATAATCAGCATCTTACGGGATGTAGA 2382

Db 2369 TGCCCAAGCAAAAGGAGGACTTATTCCTTTACATAATCGACATCTTACGGGCATGTAGA 2428  
QY 2383 TGTACGAGCTCTACTAATAAAGTATATGTCATGTGTCAATGCCAGGACAAATGGGCTTT 2442  
Db 2429 TGTAGAGCTCTACTAATAAAGTATATGTCATGTGTCAATGCCAGGACAAATGGGCTTT 2488  
QY 2443 CACACCTTTGACGAAGCAGCCAAAGGAGGACGACAGCTTTGTGCTTTGTGCTAGC 2502  
Db 2489 CACACCTTTGACGAAGCAGCCAAAGGAGGACGACAGCTTTGTGCTTTGTGCTAGC 2548  
QY 2503 CCATGGAGCTGACCGAGCTCTTAAATATCAGGAAGGACAAACACCTTTTAGATTTAGTTTC 2562  
Db 2549 CCATGGAGCTGACCGAGCTCTTAAATATCAGGAAGGACAAACACCTTTTAGATTTAGTTTC 2608  
QY 2563 AGCGATGATGTACGCGCTCTTCTGACAGCAGCCATGCCCCATCTGCTCGCCCTCTTG 2622  
Db 2609 AGCGATGATGTACGCGCTCTTCTGACAGCAGCCATGCCCCATCTGCTCGCCCTCTTG 2668  
QY 2623 TTACAAGGCTCAAGTGTCTCAATGTGTGAGAGCCAGGAGCCACTGCGAGTGTCTCTC 2682  
Db 2669 TTACAAGGCTCAAGTGTCTCAATGTGTGAGAGCCAGGAGCCACTGCGAGTGTCTCTC 2728  
QY 2683 TTCAGGTCATCTAGCCCATCAAGCCTTTCTGACGAGCAGAGCTTGTGACAACTTATCTGG 2742  
Db 2729 TTCAGGTCATCTAGCCCATCAAGCCTTTCTGACGAGCAGAGCTTGTGACAACTTATCTGG 2788  
QY 2743 GAGTTTTTTCAGAACTGTCTTCAAGTAGTTAGTTCAAGTGAACAGAGGCTTCCAGTTT 2802  
Db 2789 GAGTTTTTTCAGAACTGTCTTCAAGTAGTTAGTTCAAGTGAACAGAGGCTTCCAGTTT 2848  
QY 2803 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862  
Db 2849 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2908  
QY 2863 ACTTGAGCACCTAATGGATATATTGTGAGAGACAGATCACTTTGGATGTATTAGTTGA 2922  
Db 2909 ACTTGAGCACCTAATGGATATATTGTGAGAGACAGATCACTTTGGATGTATTAGTTGA 2968  
QY 2923 GATGGGACAAAGAGCTGAAGGAGATGGAATCAATCTTATGACATAGGACAACT 2982  
Db 2969 GATGGGACAAAGAGCTGAAGGAGATGGAATCAATCTTATGACATAGGACAACT 3028  
QY 2983 AATTAAGGAGTCGAGAGACTTATCTCGGACAAACAGGCTTTAACCATATTTAACTTT 3042  
Db 3029 AATTAAGGAGTCGAGAGACTTATCTCGGACAAACAGGCTTTAACCATATTTAACTTT 3088  
QY 3043 GAACACCTCTGTAGTGAACAAATTTCTATAGATCTGTCTCTGATGATAAGAGTTTCA 3102  
Db 3089 GAACACCTCTGTAGTGAACAAATTTCTATAGATCTGTCTCTGATGATAAGAGTTTCA 3148  
QY 3103 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTTCGAGAGCACAGATGCGAGTCTGAGG 3162  
Db 3149 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTTCGAGAGCACAGATGCGAGTCTGAGG 3208  
QY 3163 TGGAACTCTCAACAGATACAATTTCTCAAGATTCAGAGGTTTGTGAACAAGAACTATG 3222  
Db 3209 TGGAACTCTCAACAGATACAATTTCTCAAGATTCAGAGGTTTGTGAACAAGAACTATG 3268  
QY 3223 GGAAGATACACTACCGGAGAAAAGAGTTTCTGAAGAAACCAACCACTGCAATGA 3282  
Db 3269 GGAAGATACACTACCGGAGAAAAGAGTTTCTGAAGAAACCAACCACTGCAATGA 3328  
QY 3283 ACGAATGCTATTTCATGGTCTCTCTTTGTGAATCAATTTCCCAAGGCTTTGATGA 3342  
Db 3329 ACGAATGCTATTTCATGGTCTCTCTTTGTGAATCAATTTCCCAAGGCTTTGATGA 3388  
QY 3343 AAGGATGCTGCTACATAGTGTGATCTTTGGAGCTGGCATTTATTTTGTGAAAACCTCTTC 3402  
Db 3389 AAGGATGCTGCTACATAGTGTGATCTTTGGAGCTGGCATTTATTTTGTGAAAACCTCTTC 3448  
QY 3403 CAAAAGCAATCAATATGATATATGGAATTTGGAGAGGTTACTGGGTGTCAGTTTCAAAAGA 3462  
|||||

Db 3449 CAAAAGCAATCAATATATATATGGAATTTGGAGGAGGTACTGGGTCTCCAGTTCCAAAGA 3508  
QY 3463 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCGGTAACCTTTGGGAAGTC 3522  
Db 3509 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCGGTAACCTTTGGGAAGTC 3568  
QY 3523 TTTCTTGCAGTTTCAAGTGAATGAAATGACATTTCTCTCCAGGTCATCAGTCAC 3582  
Db 3569 TTTCTTGCAGTTTCAAGTGAATGAAATGACATTTCTCTCCAGGTCATCAGTCAC 3628  
QY 3583 TGGTAGGCCCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTATTACAGAGGAGAACA 3642  
Db 3629 TGGTAGGCCCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTATTACAGAGGAGAACA 3688  
QY 3643 GGCTTATCTCTGAGTATTAAATTTACAGATTTATGAGGCTTGAAGGTATGTCGATGG 3702  
Db 3689 GGCTTATCTCTGAGTATTAAATTTACAGATTTATGAGGCTTGAAGGTATGTCGATGG 3748  
QY 3703 ATAAATAGTTATTTTAAAGAACTAAATTTCCACTGAACCTAAATCATCAAGACGAGCTGG 3762  
Db 3749 ATAAATAGTTATTTTAAAGAACTAAATTTCCACTGAACCTAAATCATCAAGACGAGCTGG 3808  
QY 3763 CCTCTAGCTTTTACTCTCTTTGCTGAAAAAAA 3794  
Db 3809 CCTCTAGCTTTTACTCTCTTTGCTGAAAAAAA 3840  
|||||

RESULT 3  
AF329696 AF329696 3815 bp mRNA linear PRI 14-MAY-2001  
LOCUS Homo sapiens tankyrase 2 mRNA, complete cds.  
DEFINITION AF329696  
ACCESSION AF329696  
VERSION AF329696.1 GI:13161041  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3815)  
AUTHORS Lyons,R.J., Deane,R., Lynch,D.K., Ye,Z.S., Sanderson,G.M.,  
Eyre,H.J., Sutherland,G.R. and Daly,R.J.  
TITLE Identification of a novel human tankyrase through its interaction  
with the adaptor protein Grb14  
J. Biol. Chem. 276 (20), 17172-17180 (2001)  
MEDLINE 21264473  
PUBMED 11278563  
REFERENCE 2 (bases 1 to 3815)  
AUTHORS Lyons,R.J., Deane,R., Lynch,D.K., Ye,Z.-S.J., Sanderson,G.M.,  
Eyre,H.J., Sutherland,G.R. and Daly,R.J.  
TITLE Direct Submission  
SUBMITTED (17-DEC-2000) Cancer Research, Garvan Institute of  
Medical Research, 384 Victoria St., Sydney, NSW 2010, Australia  
FEATURES  
Location/Qualifiers  
1. 3815  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q23.2"  
140..3640  
/codon\_start=1  
/product="tankyrase 2"  
/protein\_id="AAK13463.1"  
/db\_xref="GI:13161042"  
/translation="MSGRRCAGGGAACASAAAEVPAARELFEARNGDVERVKRLV  
TPKVNSRDTAGRKSTPLHFAAGFGGRKDVVEYLLQNGANVQARDGGLIPLHNACSG  
HAEVNLRLHGDAPNARNWNTPLHEAAIKGIDVIVLQHGAEPTIRNDGRTA  
LDLADSAKAVLTGEYKDBELLESASRSGNEKMMALLTPLNVNCHASDGRKSTPLHLA  
AGYNRKIVQLLIQHGADHAKDGLVPLHNACSYGHVEYELVHKGACVNMADLW  
QFTPLHEAASKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTOLKRLAYEFKGS  
LLOAAREADVTRIKKLSLEWVNFHPOTHEALHCAASAPYKPKQICEILLRKGAN  
INKEITKEFTPLHVASSEKANDVVEVVAKEAKVNALDNLGOTSUHRAYCCHLQTCR  
LLLSYCGDPNIISLQGTFLQMGNEVQQLQEGISLGNSEADRLQLLEAAKAGDVETV  
KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRSVVEYLLQHGADVHAKGGLVPLHNA

CSYGHVEVALLVKKGVVNVADLWKFTPLHBAEAAKGYEICKLLLOHGADPTKKNRD  
GNTPLDLVKGDDTDQLDLRGAADLDAKKGCLARVKKLSSPDVNVNCRDTCGRHSTP  
LHLAAGYNLEVAEYLLQHCADVNAQDGLPLHNAASYGHWDAALLIKYNACVNA  
TDKWAFTPLHEAAQKGRTOCALLAHAGADPTLKNQEGOTPLDLVSADDSVALLTAAM  
PPSALPSCVKPVLNCSVSPGADALSSGSPSSLSAASLDNLNLSFSELSVVVS  
SSGTGASLEKEVPGVDFSTIQVRNLGLEHMDIFEREQITLDLIVEMGHKEIKE  
IGINAYGHRHKLKIGVERLISQOGLNPLYLLNTSGSGTILLDLSPDDKEFOSVEEM  
QSTVREHDDGHAGIFRNRYLKIQKVCNKKLWERYTHRRKEVEENHNHANERMLF  
HGSFPVNAIHRGFPERHAYIGGMFCAGIYFAENSXSNOYVYIGGGTGCPVHKDRS  
CVICHROLLEFCRVTLGKSPLOFSAMKMAHSPGPHSVTGRPSVNGLALAEYVYIRGEQ  
AYPEYLINQIINRPSGMDG"  
BASE COUNT 1120 a 788 c 924 g 983 t  
ORIGIN

Query Match 89.2%; Score 3387.2; DB 9; Length 3815;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	403	AGGTTTGGCGGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC	462
Db	337	AGGTTTGGCGGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAAATGTCCAAGC	396
QY	463	ACGTGATGATGGGGCTTATTCCTCTTCATAATGCATGCTCTTTGGTCATGCTGAAGT	522
Db	397	ACGTGATGATGGGGCTTATTCCTCTTCATAATGATGCTCTTTGGTCATGCTGAAGT	456
QY	523	AGTCAATCTCTTTTTCGCACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC	582
Db	457	AGTCAATCTCTTTTTCGCACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC	516
QY	583	TCCTCTCCATGAAGCTTGCAATTAAGGAAGATTTGATTTGCAATTTGCTGTGTACAGCA	642
Db	517	TCCTCTCCATGAAGCTTGCAATTAAGGAAGATTTGATTTGCAATTTGCTGTGTACAGCA	576
QY	643	TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC	702
Db	577	TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC	636
QY	703	ATCTGCCAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTCCAG	762
Db	637	ATCTGCCAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTCCAG	696
QY	763	GAGTGGCAATGAAGAAAATGATGCTCTACTCACACCATTAATTAATGTCACCTGCCACGC	822
Db	697	GAGTGGCAATGAAGAAAATGATGCTCTACTCACACCATTAATTAATGTCACCTGCCACGC	756
QY	823	AGTGATGCGAGAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	882
Db	757	AGTGATGCGAGAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	816
QY	883	TGTACAGCTGTTACTGCACATGGAGCTGATGCCATGCTAAAGATTAAGGTGATCTGGT	942
Db	817	TGTACAGCTGTTACTGCACATGGAGCTGATGCCATGCTAAAGATTAAGGTGATCTGGT	876
QY	943	ACCATTAACAATGCCGTCTTATGCTCATTATGAAGTAACCTGAACTTTTGGTCAAGCA	1002
Db	877	ACCATTAACAATGCCGTCTTATGCTCATTATGAAGTAACCTGAACTTTTGGTCAAGCA	936
QY	1003	TGGTGCTGTGTAATGCAATGGAGCTTGTGGCAATTCACCTCTTCATCAGGCAGCTTC	1062
Db	937	TGGTGCTGTGTAATGCAATGGAGCTTGTGGCAATTCACCTCTTCATCAGGCAGCTTC	996
QY	1063	TGAAGACGGTGAAGTATGTTCTCTTCTTAAGTTATGTTGCGAGACCCCAACACTGCT	1122
Db	997	TGAAGACGGTGAAGTATGTTCTCTTCTTAAGTTATGTTGCGAGACCCCAACACTGCT	1056
QY	1123	CAATTGTCAATAAAAGTGCATAGACTTGGCTCCACACACACAGTTAAAAGAAAAGATT	1182
Db	1057	CAATTGTCAATAAAAGTGCATAGACTTGGCTCCACACACACAGTTAAAAGAAAAGATT	1116
QY	1183	AGCATATGAATTTAAAGGCCACTCGTGTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG	1242
Db	1117	AGCATATGAATTTAAAGGCCACTCGTGTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG	1176

QY	1243	AATCAAAAACATCTCTCTCTGGAAATGGTGAATTTTCAAGCATCTCTCAAAACACATGAAC	1302
Db	1177	AATCAAAAACATCTCTCTCTGGAAATGGTGAATTTTCAAGCATCTCTCAAAACACATGAAC	1236
QY	1303	AGCATTTGCATTTGTGCTGCATCTCCATATCCCAAAAGAAACAAATATGTGAACATGTT	1362
Db	1237	AGCATTTGCATTTGTGCTGCATCTCCATATCCCAAAAGAAACAAATATGTGAACATGTT	1296
QY	1363	GCTAAGAAAGGAGCAACACATCAATGAAAAGACTTAAGAAATCTTGTGCTCCTCTGCACGT	1422
Db	1297	GCTAAGAAAGGAGCAACACATCAATGAAAAGACTTAAGAAATCTTGTGCTCCTCTGCACGT	1356
QY	1423	GGCATCTGAGAAAAGCTCATTAATGATTTTGAAGTAGTGGTGAACATCAAGCAAAAGT	1482
Db	1357	GGCATCTGAGAAAAGCTCATTAATGATTTTGAAGTAGTGGTGAACATCAAGCAAAAGT	1416
QY	1483	TAATGCTGTGATAATCTTGGTCAGACTTCTCTACACAGAGCTGTCATATTTGTTGCTCATCT	1542
Db	1417	TAATGCTGTGATAATCTTGGTCAGACTTCTCTACACAGAGCTGTCATATTTGTTGCTCATCT	1476
QY	1543	ACAAACCTGCCGCTACTCCTGAGCTATGGTGTGATCCCTAAACATATATCCCTTCAGGG	1602
Db	1477	ACAAACCTGCCGCTACTCCTGAGCTATGGTGTGATCCCTAAACATATATCCCTTCAGGG	1536
QY	1603	CTTTACTGCTTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTC	1662
Db	1537	CTTTACTGCTTTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTC	1596
QY	1663	ATTAGTAATTCAGAGGCACAGACAAATTTGCTGGAAGCTGCAAGAGCTGCAATGCTCA	1722
Db	1597	ATTAGTAATTCAGAGGCACAGACAAATTTGCTGGAAGCTGCAAGAGCTGCAATGCTCA	1656
QY	1723	AACGTGAAAAAACTGTACTGTTTACAGAGTGTCAACTGCAGAGACATTTCAAGGGGTCA	1782
Db	1657	AACGTGAAAAAACTGTACTGTTTACAGAGTGTCAACTGCAGAGACATTTCAAGGGGTCA	1716
QY	1783	GTCACACACTTTCATTTTGCAGCTGGGTATAACAGAGTGTGCTGGTGGGAATATCTGCT	1842
Db	1717	GTCACACACTTTCATTTTGCAGCTGGGTATAACAGAGTGTGCTGGTGGGAATATCTGCT	1776
QY	1843	ACAGCATGAGCTGATGTCATGCTTAAAGATAAGAGGCGCTTGTACCTTTGCAACATGC	1902
Db	1777	ACAGCATGAGCTGATGTCATGCTTAAAGATAAGAGGCGCTTGTACCTTTGCAACATGC	1836
QY	1903	ATGTTCTTATGACATATTAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTAA	1962
Db	1837	ATGTTCTTATGACATATTAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTAA	1896
QY	1963	TGTAGCTGATTTATGAAATTTTACCTTTTACATGAAGCAGCAGCAAAAGGAAAATATGA	2022
Db	1897	TGTAGCTGATTTATGAAATTTTACCTTTTACATGAAGCAGCAGCAAAAGGAAAATATGA	1956
QY	2023	AATTTCCAAACTTCTGCTCAGCATGTCAGACCTTACCACAAAAAAGAGGATGAAA	2082
Db	1957	AATTTCCAAACTTCTGCTCAGCATGTCAGACCTTACCACAAAAAAGAGGATGAAA	2016
QY	2083	TACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA	2142
Db	2017	TACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA	2076
QY	2143	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAGTTGCTTTC	2202
Db	2077	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAGTTGCTTTC	2136
QY	2203	TCTGTATAATGTAATTTGCCGGGATACCACAGGCACACATTTCAACACCTTTTACATTTAGC	2262
Db	2137	TCTGTATAATGTAATTTGCCGGGATACCACAGGCACACATTTCAACACCTTTTACATTTAGC	2196
QY	2263	AGCTGTTTATAATTAATTTAGAAAGTTCAGAGATATTTTGTACACACGAGCTGATGTGAA	2322
Db	2197	AGCTGTTTATAATTAATTTAGAAAGTTCAGAGATATTTTGTACACACGAGCTGATGTGAA	2256



GEQAYPEVLYTYQIMRPEGVMDG"

3509...4127

1245 a 813 c 934 g 1135 t

3'UTR

BASE COUNT

ORIGIN

Query Match 89.2%; Score 3387.2; DB 6; Length 4127;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3389; Conservative 0; Mismatches

0;

Gaps

Indels

0;

Gaps

QY 403 AGTTTGGCGGGAAGACGTAGTTGAATATTTGCTTCAGAAATGTTGCAAGTGTCCAAGC 462

Db 208 AGTTTGGCGGGAAGACGTAGTTGAATATTTGCTTCAGAAATGTTGCAAGTGTCCAAGC 267

QY 463 AGCTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 522

Db 268 AGCTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 327

QY 523 AGTCAATCTCCTTTTTCGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATATATAC 582

Db 328 AGTCAATCTCCTTTTTCGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATATATAC 387

QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTTGATGTTTGCATTTGCTGTTCACAGCA 642

Db 388 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTTGATGTTTGCATTTGCTGTTCACAGCA 447

QY 643 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTTAGCAGATCC 702

Db 448 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTTAGCAGATCC 507

QY 703 ATCTGCGAAGACGTGCTTACTGCTGAATATAAGAAAGATGAACCTCTTAGAAAGTCCAG 762

Db 508 ATCTGCGAAGACGTGCTTACTGCTGAATATAAGAAAGATGAACCTCTTAGAAAGTCCAG 567

QY 763 GAGTGGCAATGAAGAAATGATGGCTCTACTCACACCAATTAATGTCAACTGCCACGC 822

Db 568 GAGTGGCAATGAAGAAATGATGGCTCTACTCACACCAATTAATGTCAACTGCCACGC 627

QY 823 AAGTGATGGCAGAAAGTCACTCCATTTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882

Db 628 AAGTGATGGCAGAAAGTCACTCCATTTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

QY 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCATTAAGATTAAGAGTATGATCTGGT 942

Db 688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCATTAAGATTAAGAGTATGATCTGGT 747

QY 943 ACCATTTACAAATGCCCTTCTTATGCTCATTTAAGTCACTGCTTAAAGATTAAGAGTATG 1002

Db 748 ACCATTTACAAATGCCCTTCTTATGCTCATTTAAGTCACTGCTTAAAGATTAAGAGTATG 807

QY 1003 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGCAGCTTC 1062

Db 808 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGCAGCTTC 867

QY 1063 TAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTATGAGTATGAGTATGAGTATGAGTATG 1122

Db 868 TAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTATGAGTATGAGTATGAGTATGAGTATG 927

QY 1123 CAATTTGCACAAATAAAGTGTATAGACTTGGCTCCACACACACAGTTTAAAGAAAGATT 1182

Db 928 CAATTTGCACAAATAAAGTGTATAGACTTGGCTCCACACACACAGTTTAAAGAAAGATT 987

QY 1183 AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCACGAGAAGCTGATTTACTTCG 1242

Db 988 AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCACGAGAAGCTGATTTACTTCG 1047

QY 1243 AATCAAAAACATCTCTCTTGAAATGGTGAATTTCAAGCATCTTCAACACATGAAC 1302

Db 1048 AATCAAAAACATCTCTCTTGAAATGGTGAATTTCAAGCATCTTCAACACATGAAC 1107

QY 1303 AGCATTTGATTTGCTGCTGATCTCCATATCCCAAAAAGAAAGCAATATGTGAATCTTT 1362

Db 1108 AGCATTTGATTTGCTGCTGATCTCCATATCCCAAAAAGAAAGCAATATGTGAATCTTT 1167

QY 1363 GCTAAGAAAGGAGCAACATCAATGAAAGACTAAAGAAATTTGACCTCTCTGCACGT 1422

Db 1168 GCTAAGAAAGGAGCAACATCAATGAAAGACTAAAGAAATTTGACCTCTCTGCACGT 1227

QY 1423 GGCATCTGAGAAGCTCATAATGATGTTGTTGAAAGTAGTGGTGAACATCAAGCAAGGT 1482

Db 1228 GGCATCTGAGAAGCTCATAATGATGTTGTTGAAAGTAGTGGTGAACATCAAGCAAGGT 1287

QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCAATATTTGGTTCATCT 1542

Db 1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCAATATTTGGTTCATCT 1347

QY 1543 ACAAACTGCCCTACTCTGAGCTATGGGTGATGCTTAACATTTATATCCCTTCAGG 1602

Db 1348 ACAAACTGCCCTACTCTGAGCTATGGGTGATGCTTAACATTTATATCCCTTCAGG 1407

QY 1603 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGAACTCTTCAAGAGGATATCTC 1662

Db 1408 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGAACTCTTCAAGAGGATATCTC 1467

QY 1663 ATTAGTAAATTCAGAGGCAGACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCTGA 1722

Db 1468 ATTAGTAAATTCAGAGGCAGACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCTGA 1527

QY 1723 AACTGTAAAAAACTGTGTACTCTTCCAGAGTGTCAACTGCAGAGACATTTCAAGGGCTCA 1782

Db 1528 AACTGTAAAAAACTGTGTACTCTTCCAGAGTGTCAACTGCAGAGACATTTCAAGGGCTCA 1587

QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTCTCGTGGTGAATATCTGCT 1842

Db 1588 GTCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTCTCGTGGTGAATATCTGCT 1647

QY 1843 ACAGCATGGAGCTGATGTGCATCTTAAAGATAAAGGAGGCTTGTACCTTTGCAACAATGC 1902

Db 1648 ACAGCATGGAGCTGATGTGCATCTTAAAGATAAAGGAGGCTTGTACCTTTGCAACAATGC 1707

QY 1903 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTCTTGTAAACATGAGCAGTAGTTAA 1962

Db 1708 ATGTTCTTATGGACATTTATGAAGTTGCAGAACTCTTGTAAACATGAGCAGTAGTTAA 1767

QY 1963 TGTAGCTGATTTTGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA 2022

Db 1768 TGTAGCTGATTTTGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA 1827

QY 2023 AATTTGCAAACTTCTGCTCCAGCATGTTGCAGACCCCTACCAAAAAACAGGGATGAAA 2082

Db 1828 AATTTGCAAACTTCTGCTCCAGCATGTTGCAGACCCCTACCAAAAAACAGGGATGAAA 1887

QY 2083 TACTCCTTTGATCTTCTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGAGA 2142

Db 1888 TACTCCTTTGATCTTCTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGAGA 1947

QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAGTTGCTPTC 2202

Db 1948 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAGTTGCTPTC 2007

QY 2203 TCCGTGATAAGTAAATTTGCCCGCATACCCAGGCGAGACATTTCAACACCTTTACATTTAGC 2262

Db 2008 TCCGTGATAAGTAAATTTGCCCGCATACCCAGGCGAGACATTTCAACACCTTTACATTTAGC 2067

QY 2263 AGCTGGTTATAAATTTAGAGTTTGCAGAGTATTGTTTACAACACGAGCTGATGTGAA 2322

Db 2068 AGCTGGTTATAAATTTAGAGTTTGCAGAGTATTGTTTACAACACGAGCTGATGTGAA 2127

QY 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGAGCATTTACGGGCATGTAGA 2382

Db 2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGAGCATTTACGGGCATGTAGA 2187

QY 2383 TGTAGCAGCTCTACTAATAAAGTATTAATGCAATGCTGCTCAATGCGCAGCAAAATGGCTTT 2442

Db 2188 TGTAGCAGCTCTACTAATAAAGTATTAATGCAATGCTGCTCAATGCGCAGCAAAATGGCTTT 2247

QY 2443 CACACCTTTGCGAAGCAGCAGCCCAAGGAGGAGCAACACAGCTTTTGTGCTTAGC 2502

|||||  
Db 2248 CACACCTTTGCAGAGCACCCAAAAGGACCAACACAGCTTTGTGCTTTGTGTAGC 2307  
QY 2503 CCATGGAGCTGACCCGACTCTTAAAATCAGGAAGACAAACACCTTTAGATTAGTTTC 2562  
Db 2308 CCATGGAGCTGACCCGACTCTTAAAATCAGGAAGACAAACACCTTTAGATTAGTTTC 2367  
QY 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTGCGCCCTTG 2622  
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTGCGCCCTTG 2427  
QY 2623 TTACAAGCCTCAAGTGTCTCAATGCTGAGAGCCAGAGCCAGAGCCAGAGTCTCTCTC 2682  
Db 2428 TTACAAGCCTCAAGTGTCTCAATGCTGAGAGCCAGAGCCAGAGTCTCTCTC 2487  
QY 2683 TTCAAGTCTCATAGCCCATCAAGCCCTTTCTGAGAGCCAGAGTCTTGTGACAACTTATCTGG 2742  
Db 2488 TTCAAGTCTCATAGCCCATCAAGCCCTTTCTGAGAGCCAGAGTCTTGTGACAACTTATCTGG 2547  
QY 2743 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2802  
Db 2548 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2607  
QY 2803 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCGTAAGGAATCTTGG 2862  
Db 2608 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCGTAAGGAATCTTGG 2667  
QY 2863 ACTTGAGCCTCAATGGATATATTTTGAGAGAACAGATCACTTTGGATGTTAGTTGA 2922  
Db 2668 ACTTGAGCCTCAATGGATATATTTTGAGAGAACAGATCACTTTGGATGTTAGTTGA 2727  
QY 2923 GATGGGCAACAGGAGCTGAAGAGATTTGGAATCAATGCTTATGACATAGGCACAACT 2982  
Db 2728 GATGGGCAACAGGAGCTGAAGAGATTTGGAATCAATGCTTATGACATAGGCACAACT 2787  
QY 2983 AATTAAGGAGTCCAGAGACTTATCTCCGACAAAGGCTTAACCCATATTAACTTT 3042  
Db 2788 AATTAAGGAGTCCAGAGACTTATCTCCGACAAAGGCTTAACCCATATTAACTTT 2847  
QY 3043 GAACACCTCTGGTAGTGAACAATTTTATAGATCTGCTCTGATGATGATAAGAGTTTCA 3102  
Db 2848 GAACACCTCTGGTAGTGAACAATTTTATAGATCTGCTCTGATGATAAGAGTTTCA 2907  
QY 3103 GTCTGTGGAGGAGAGATCCAAAGTACAGTTTCAGAGTTCAGAGCAGACAGATGAGGTCATGCAGG 3162  
Db 2908 GTCTGTGGAGGAGAGATCCAAAGTACAGTTTCAGAGTTCAGAGCAGACAGATGAGGTCATGCAGG 2967  
QY 3163 TGGAACTCTTCAACAGATACAATATTCTCAAGATTTCAGAGGTTTCTGAACAAGAACTATG 3222  
Db 2968 TGGAACTCTTCAACAGATACAATATTCTCAAGATTTCAGAGGTTTCTGAACAAGAACTATG 3027  
QY 3223 GGAAGATACACTCACCGGAGAAAAGGTTTCTGAAGAAAACCAACCAATGCCAATGA 3282  
Db 3028 GGAAGATACACTCACCGGAGAAAAGGTTTCTGAAGAAAACCAACCAATGCCAATGA 3087  
QY 3283 AGCAATGCTATTTCATGGGTCCTTTTGTGATGATCTGCTCTGATGATGATAAGAGTTTCA 3342  
Db 3088 AGCAATGCTATTTCATGGGTCCTTTTGTGATGATCTGCTCTGATGATAAGAGTTTCA 3147  
QY 3343 AAGGATGCTGATACAGTGGTATGTTTGGAGTGGCATTTATTTGCTGAAACTCTTC 3402  
Db 3148 AAGGATGCTGATACAGTGGTATGTTTGGAGTGGCATTTATTTGCTGAAACTCTTC 3207  
QY 3403 CAAAGCAATCAATATGATATGGAATTTGGAGAGGTAAGTGGGTCCTGATGATGATAAGAG 3462  
Db 3208 CAAAGCAATCAATATGATATGGAATTTGGAGAGGTAAGTGGGTCCTGATGATAAGAG 3267  
QY 3463 CAGATCTGTTACATTTGCCAGAGCAGTGTCTTTTCCGCGGTAACCTTTGGGAAAGTC 3522  
Db 3268 CAGATCTGTTACATTTGCCAGAGCAGTGTCTTTTCCGCGGTAACCTTTGGGAAAGTC 3327  
QY 3523 TTCTCTGAGTTTCAGTGAATGAAATGGACATCTCTCTCCAGTTCATCACTCAGTCAC 3582  
|||||

Db 3328 TTCTCTGCACTTTCAGTGCATGCAATGAAATGGACATTTCTCTCCAGGTCATCACTCAGTCAC 3387  
QY 3583 TGGTAGGCCAGCTGTAATGGCTAGCATTTAGCTGAATATGTTATTTTACAGAGGAGACA 3642  
Db 3388 TGGTAGGCCAGCTGTAATGGCTAGCATTTAGCTGAATATGTTATTTTACAGAGGAGACA 3447  
QY 3643 GCTTATCTCTGAGTATTTAAATTTACTTACCAGATTTAGAGGCTGAAGTATGTTGTCATGG 3702  
Db 3448 GCTTATCTCTGAGTATTTAAATTTACTTACCAGATTTAGAGGCTGAAGTATGTTGTCATGG 3507  
QY 3703 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGG 3762  
Db 3508 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGG 3567  
QY 3763 CCTCTACGTTTCTACTCCTTTGCTGAAAAAAA 3794  
Db 3568 CCTCTACGTTTCTACTCCTTTGCTGAAAAAAA 3599  
RESULT 5  
AX062247  
LOCUS AX062247 4406 bp DNA linear PAT 24-JAN-2001  
DEFINITION Sequence 106 from Patent WO0100849.  
ACCESSION AX062247  
VERSION AX062247.1 GI:12540148  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4406)  
AUTHORS Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.  
TITLE Tankyrase2 materials and methods  
JOURNAL Patent: WO 0100849-A 106 04-JAN-2001;  
ICOS CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..4406  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
<2..3790  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC25062.1"  
/db\_xref="GI:12540149"  
/translation="ELAGGALPASAAASFODDPGFALPPPRSGAGSPARGARG  
GHGTADPPTAGSQAALSSAPSGGLALLAGPGLLLRLLALLLAAARINSGRR  
AGGACASAAEAPEARELFEACRNGDVERVKRLVTPKVNVRDTRAGKSTPLHF  
AAGFRKDVVEYLLQNGANVOARDGGLIPLHNACSFHAEVNVLLLRHGDNFARD  
WNTPLHEAA IKGI DVCIVLQHGAEPTIRNTDRTALDADPSAKAVLTGYKDE  
LLESARGNEEMMALITPLNVNCHASDGRKSTPLAAGYNRVKIVQLLQHGADVH  
AKDGLVPLHNACSYGHVEYVELLKHGACVNMALWOFTPLHEAASKNRVVCSSL  
LSYGADPTLLNCHNSAIDLAPTPOLKERLAYEFKSHLSLOAREADYTRIKHLSLE  
MYNFKPOTHEALTHKNSAIDLPYKQICELLURKANINEKTEFLPLHVASEKAL  
NDVEVVVHEAKNALDNIGQTSLHRAAYCGHLQTRLLLSYCDNPILISQGETAL  
OMGENVQQLQEGISLGNSEADROLLEAAKAGDVETVKLCTQSVNCRDIEGROST  
PLHFAAGYNRVSYVEYLLQHGADVHAKDGLVPLHNACSYGHVEYVELLQHGAVN  
VADLWFTPLHEAAKGYEICKLLQHGADPTKKNRDNTPDLIVKDGDTDIDLRL  
GDAALLDAAKKGLARVKLSLSPDYNCRDTQGRHSTPLHLAAGYNLEVEYLLQHG  
ADVNAQDKGLIPLHNAASGYHVDVAALLIKYNACVNDATDKWAFPLHEAAQGRTO  
CALLAHAGADPTLNQEGOTPLDLVSADVSALLTAAMPSPALPSYKPOVLNGVRS  
GATADALSPPSGSPSSLSAASLNDLSGSFSELSVSVSSSGTSGASLSEKKEVGVDF  
STQFVFNGLHLEHMDIFEREQITLDVLVEMGHKEIGINAYHRHKLKGVRLI  
SGOGLNPLYLTNTSGSTILIDLSPPDKFQSFVEEMOSTVREHEDGCHAGINRY  
NILKIOKNKLLWERYTHRRKVESENHNANERMLFHGSPFVNAIHKGFDERHAY  
IGMFCAGIYFAENSNSNOYVTYIGGTCGCPVHKDRSCYICHRLLFCRVLTGKSLF  
QPSAMHAHSPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLYTIQIMRPEGWVG  
"

BASE COUNT 1268 a 910 c 1054 g 1174 t  
ORIGIN

Query Match 89.2%; Score 3387.2; DB 6; Length 4406;  
Best Local Similarity 99.9%; Pred. 0.0;  
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;



QY	403	AGGTTTTGGCGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC	462	QY	1483	TAATGCTTCGGATAAATCTTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCT	1542
Db	487	AGGTTTTGGCGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC	546	Db	1567	TAATGCTTCGGATAAATCTTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCT	1626
QY	463	AGTGATGATGGGGCCTTATTCCTCTTCATATATGATGCTCTTTTGGTCACTGCTGAAGT	522	QY	1543	ACAAACCTCCGCCCTACTCTCTGAGCTATGGGTGATGCTTAACATTTATATCCCTTCAGGG	1602
Db	547	AGTGATGATGGGGCCTTATTCCTCTTCATATATGATGCTCTTTTGGTCACTGCTGAAGT	606	Db	1627	ACAAACCTCCGCCCTACTCTCTGAGCTATGGGTGATGCTTAACATTTATATCCCTTCAGGG	1686
QY	523	AGTCAATCTCCTTTTCGGACATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATATATC	582	QY	1603	CTTTACTGCTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTC	1662
Db	607	AGTCAATCTCCTTTTCGGACATGGTGACAGACCCCAATGCTCGAGATAATTTGGAATATATC	666	Db	1687	CTTTACTGCTTACAGATGGAAATGAAATGTACAGCAACTCTCTCAAGAGGCTATCTC	1746
QY	583	TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTGTCATTGCTGTTCACAGCA	642	QY	1663	ATTAGGTAATTCAGAGGCACAGACAAATTTGCTGGAAGTGCAGAAAGCTGGAGATGCTGA	1722
Db	667	TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTGTCATTGCTGTTCACAGCA	726	Db	1747	ATTAGGTAATTCAGAGGCACAGACAAATTTGCTGGAAGTGCAGAAAGCTGGAGATGCTGA	1806
QY	643	TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCAATTTGATTTAGCAGATCC	702	QY	1723	AACTGTAAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTGCAAGGGGTCA	1782
Db	727	TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCAATTTGATTTAGCAGATCC	786	Db	1807	AACTGTAAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTGCAAGGGGTCA	1866
QY	703	ATCTGCCAAGCAGTCTTACTGGTGAATATAGAAAGATGAACCTCTTAGAAAGTCCAG	762	QY	1783	GTCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT	1842
Db	787	ATCTGCCAAGCAGTCTTACTGGTGAATATAGAAAGATGAACCTCTTAGAAAGTCCAG	846	Db	1867	GTCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT	1926
QY	763	GAGTGCAATGAAGAAAAATGATGCTCTACTCACACCATTAATGTCAACTGCCACGC	822	QY	1843	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTTGTACCTTTGCACATGC	1902
Db	847	GAGTGCAATGAAGAAAAATGATGCTCTACTCACACCATTAATGTCAACTGCCACGC	906	Db	1927	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTTGTACCTTTGCACATGC	1986
QY	823	AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	882	QY	1903	ATGTTCTTATGGACATTTACAACTTCAGAACTTCTGTTTAAACATGGAGCAGTAGTTAA	1962
Db	907	AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	966	Db	1987	ATGTTCTTATGGACATTTACAACTTCAGAACTTCTGTTTAAACATGGAGCAGTAGTTAA	2046
QY	883	TGTACAGCTTTACTGCACATGGAGCTGATGTCATGCTTAAAGATAAAGGATCTGGT	942	QY	1963	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2022
Db	967	TGTACAGCTTTACTGCACATGGAGCTGATGTCATGCTTAAAGATAAAGGATCTGGT	1026	Db	2047	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2106
QY	943	ACCATACACAATGCGCTTGTATGCTTATGATGATTAAGTAACTGCACTTTTGGTCAAGCA	1002	QY	2023	AATTTGCAAACTTCTGCTCCAGCATGGTGAGACCCCTACCAAAAAACAGGATGGA	2082
Db	1027	ACCATACACAATGCGCTTGTATGCTTATGATGATTAAGTAACTGCACTTTTGGTCAAGCA	1086	Db	2107	AATTTGCAAACTTCTGCTCCAGCATGGTGAGACCCCTACCAAAAAACAGGATGGA	2166
QY	1003	TGTTGCTGTGTAATGCAATGGACTTGGGCAATTCACCTCTTCAATGAGGCAGCTTC	1062	QY	2083	TACTCTTTGATCTGTTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA	2142
Db	1087	TGTTGCTGTGTAATGCAATGGACTTGGGCAATTCACCTCTTCAATGAGGCAGCTTC	1146	Db	2167	TACTCTTTGATCTGTTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA	2226
QY	1063	TAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGTTGTCAGACCCCAACTGCT	1122	QY	2143	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAAGTTGCTTC	2202
Db	1147	TAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGTTGTCAGACCCCAACTGCT	1206	Db	2227	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAAGTTGCTTC	2286
QY	1123	CAATTGTCACAAATAAAGTGTCTATAGACTTGGCTCCACACCACAGTAAAGAAAGATT	1182	QY	2203	TCCTGATAATGTAATTTGCCCGCATACCCAGGCAGACATTCAACACCTTTACATTTAGC	2262
Db	1207	CAATTGTCACAAATAAAGTGTCTATAGACTTGGCTCCACACCACAGTAAAGAAAGATT	1266	Db	2287	TCCTGATAATGTAATTTGCCCGCATACCCAGGCAGACATTCAACACCTTTACATTTAGC	2346
QY	1183	AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG	1242	QY	2263	AGCTGGTTATAATAATTTAGAAAGTTGCAGAGTATTTCTTACAACACAGAGTGAATGAA	2322
Db	1267	AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG	1326	Db	2347	AGCTGGTTATAATAATTTAGAAAGTTGCAGAGTATTTCTTACAACACAGAGTGAATGAA	2406
QY	1243	AATCAAAAAACATCTCTCTCTGAAATGGTGAATTTCAAGCATCTCTCAAAACACATGAAC	1302	QY	2323	TGCCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGTCAGCATCTTACGGGCATGAG	2382
Db	1327	AATCAAAAAACATCTCTCTCTGAAATGGTGAATTTCAAGCATCTCTCAAAACACATGAAC	1386	Db	2407	TGCCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGTCAGCATCTTACGGGCATGAG	2466
QY	1303	AGCATGCAATTTGCTGCTGCATCTCCATATCCCATATCCCAAAAGAAAGCAATATGTAAC	1362	QY	2383	TGTAGCAGCTTACTATAAAGTATATGATGTGTCAATGCCACGACAAATGGGCTTT	2442
Db	1387	AGCATGCAATTTGCTGCTGCATCTCCATATCCCATATCCCAAAAGAAAGCAATATGTAAC	1446	Db	2467	TGTAGCAGCTTACTATAAAGTATATGATGTGTCAATGCCACGACAAATGGGCTTT	2526
QY	1363	GCTAAGAAAAGGACCAACATCAATGAAGACTTAAGAAATCTTGTGCTCTCTGACGCT	1422	QY	2443	CACACCTTTTCAGCAAGCAGCCCAAAAGGAGGACGAACACAGCTTTGTGCTTTAGTTC	2502
Db	1447	GCTAAGAAAAGGACCAACATCAATGAAGACTTAAGAAATCTTGTGACTCTCTGACGCT	1506	Db	2527	CACACCTTTTCAGCAAGCAGCCCAAAAGGAGGACGAACACAGCTTTGTGCTTTAGTTC	2586
QY	1423	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGAAGCAATGAAGCAAAAGGT	1482	QY	2503	CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC	2562
Db	1507	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGAAGCAATGAAGCAAAAGGT	1566	Db	2587	CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC	2646
				QY	2563	AGCGGATGATGTACAGCCTCTTCTGTACAGGAGCCATGCCCCACTCTGCTGCTGCTCTTG	2622





D	b	1193	AGTCAATCTCCTTTTGGCATGGTGCAGCCCCAATGCTCGAGATAATTGGAAATTATAC	1251
Q	y	583	TCCTCTCCATCAAGCTGCAATTTAAAGGAAGATTCATGTTTGCATTGTGCTGTTTACAGCA	642
D	b	1253	TCCTCTCCATGAAGCTGCAATTAAGGAAGATTCATGTTTGCATTGTGCTGTTTACAGCA	1312
Q	y	643	TGGAGCTGAGCCAAACCATCCGAATACAGATGGAAGGACAGCATTTGATGATTTAGCAGATCC	702
D	b	1313	TGGAGCTGAGCCAAACCATCCGAATACAGATGGAAGGACAGCATTTGATGATTTAGCAGATCC	1372
Q	y	703	ATCTGCCAAAAGCAGTCTTACTGGTGAATATAGAAAGATGAACCTCTTTAGAAAGTGCAG	762
D	b	1373	ATCTGCCAAAGCAGTCTTACTGGTGAATATAGAAAGATGAACCTCTTAGAAAGTGCAG	1432
Q	y	763	GAGTGGCAATTGAAGAAAAATGATGCTCTACTCACACCATTAATATGTCAACTGCCACGC	822
D	b	1433	GAGTGGCAATTGAAGAAAAATGATGCTCTACTCACACCATTAATATGTCAACTGCCACGC	1492
Q	y	823	AAGTGATGGCAGAAAGTCAACTCCATTACATTTGSCACGAGATATACAGAGTAAGAAT	882
D	b	1493	AAGTGATGGCAGAAAGTCAACTCCATTACATTTGSCACGAGATATACAGAGTAAGAAT	1552
Q	y	883	TGTACAGCTGTTACTCCAACTGGAGCTGATGCTCCATGCTTAAAGATAAAGTGATCTGGT	942
D	b	1553	TGTACAGCTGTTACTCCAACTGGAGCTGATGCTCCATGCTTAAAGATAAAGTGATCTGGT	1612
Q	y	943	ACCATTACAAATGCCGTCTTCTATTGGTCATTATCAAGTAAGTGAACCTTTGGTCAAGCA	1002
D	b	1613	ACCATTACAAATGCCGTCTTCTATTGGTCATTATCAAGTAAGTGAACCTTTGGTCAAGCA	1672
Q	y	1003	TGTCCTGCTGTTAAATGCAATGCACTTGTGGCAATTCACCTCTTCATGAGCAGCTTC	1062
D	b	1673	TGTCCTGCTGTTAAATGCAATGCACTTGTGGCAATTCACCTCTTCATGAGCAGCTTC	1732
Q	y	1063	TAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAGTTATGGTGCAGACCCACACGTGCT	1122
D	b	1733	TAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAGTTATGGTGCAGACCCACACGTGCT	1792
Q	y	1123	CAATTGTCACAATAAAAGTGCTATAGACTGGCTCCACACACACAGTTTAAAAGAAAGATT	1182
D	b	1793	CAATTGTCACAATAAAAGTGCTATAGACTGGCTCCACACACACAGTTTAAAAGAAAGATT	1852
Q	y	1183	AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG	1242
D	b	1853	AGCATATGAATTTAAAGGCCACTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTCG	1912
Q	y	1243	AATCAAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCATCCTCAAAACACATGAAC	1302
D	b	1913	AATCAAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCATCCTCAAAACACATGAAC	1972
Q	y	1303	AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAAAAGAACAAATATGTGAACGTGTT	1362
D	b	1973	AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAAAAGAACAAATATGTGAACGTGTT	2032
Q	y	1363	GCTAAGAAAAGGAGCAACACATCAATGAAAGACTTAAGAAATTCCTTGACTCCTCTGCACGT	1422
D	b	2033	GCTAAGAAAAGGAGCAACACATCAATGAAAGACTTAAGAAATTCCTTGACTCCTCTGCACGT	2092
Q	y	1423	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAAGACATGAAGCAAAAGT	1482
D	b	2093	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAAGACATGAAGCAAAAGT	2152
Q	y	1483	TAATGCTCTGGATAATCTTGGTGCACACTCTCTACACAGAGCTGCATTTGTGGTCACT	1542
D	b	2153	TAATGCTCTGGATAATCTTGGTGCACACTCTCTACACAGAGCTGCATTTGTGGTCACT	2212
Q	y	1543	ACAAAGCTGCGGCTACTCCTCAGCTATGGGTGTGATCCTTAACATATATCCCTTCAGGG	1602
D	b	2213	ACAAAGCTGCGGCTACTCCTCAGCTATGGGTGTGATCCTTAACATATATCCCTTCAGGG	2272
Q	y	1603	CTTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCACTCCTCAAGAGGATCTCTC	1662

Db	2273	CTTTACTGCTTTACAGATTGGGAAATGAAAATGTACAGCAACTCCTCTCAAGAGGGTATCTC	2333
Qy	1663	ATTAGCTAATTACAGAGCAGACAGACAATTTGCTGGAGCTGCCAAAGGCTCGAGATGTCGA	1732
Db	2333	ATTAGCTAATTACAGAGCAGACAGACAATTTGCTGGAGCTGCCAAAGGCTCGAGATGTCGA	2392
Qy	1723	AACTGTAAAAAACTGTGTACTGTTCCAGAGTGCAACTGCAGAGACANTTGAAGGGCGTCA	1782
Db	2393	AACTGTAAAAAACTGTGTACTGTTCCAGAGTGCAACTGCAGAGACANTTGAAGGGCGTCA	2452
Qy	1783	GTCTACACCACCTTCATTTTGCAGCTGGGTATAACAGAGTCTCGTGGTGAATATCTGCT	1842
Db	2453	GTCTACACCACCTTCATTTTGCAGCTGGGTATAACAGAGTCTCGTGGTGAATATCTGCT	2512
Qy	1843	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTTGTACCTTTGCACAAATGC	1902
Db	2513	ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTTGTACCTTTGCACAAATGC	2572
Qy	1903	ATGTTCTTATGGACATTATGAAGCTTCAGAACTCTTGTTTAAACATGGAGCAGTACGTAA	1962
Db	2573	ATGTTCTTATGGACATTATGAAGCTTCAGAACTCTTGTTTAAACATGGAGCAGTACGTAA	2632
Qy	1963	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2022
Db	2633	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2692
Qy	2023	AATTTGCAAACTTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAACAGGATGGAAA	2082
Db	2693	AATTTGCAAACTTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAACAGGATGGAAA	2752
Qy	2083	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA	2142
Db	2753	TACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA	2812
Qy	2143	TGCAGCTTTGCTAGATGCTGCCAAGAAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTTC	2202
Db	2813	TGCAGCTTTGCTAGATGCTGCCAAGAAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTTC	2872
Qy	2203	TCTGTATATGTAATTTGCCGCGATACCAAGGCAGACATTTCAACACCTTTACATTTAGC	2262
Db	2873	TCTGTATATGTAATTTGCCGCGATACCCAGCCACATTTCAACACCTTTACATTTAGC	2932
Qy	2263	AGCTGGTTATAATTTAGAAGTTCAGAGTATTTGTTTACACACGGAGCTGATGTGAA	2322
Db	2933	AGCTGGTTATAATTTAGAAGTTCAGAGTATTTGTTTACACACGGAGCTGATGTGAA	2992
Qy	2323	TGCCCCAAGCAAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGATGTAGA	2382
Db	2993	TGCCCCAAGCAAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGATGTAGA	3052
Qy	2383	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT	2442
Db	3053	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT	3112
Qy	2443	CACACCTTTGCAGAGCAGCCCCAAAAGGGACAAACACAGCTTTGTGCTTTGCTAGC	2502
Db	3113	CACACCTTTGCAGAGCAGCCCCAAAAGGGACAAACACAGCTTTGTGCTTTGTTGCTAGC	3172
Qy	2503	CCATGGAGCTGACCCGACCTCTTTAAAAATCAGGAAGCAGCAAAACCTTTAGATTTAGTTTC	2562
Db	3173	CCATGGAGCTGACCCGACCTCTTTAAAAATCAGGAAGCAGCAAAACCTTTAGATTTAGTTTC	3232
Qy	2563	ACGGAGATGTCAGGGCTCTTCTTGACAGAGCCATGCCGCCCAATGTGCTTGCCCTCTTG	2622
Db	3233	ACGAGATGATGTACGGCTCTTCTTGACAGCAGCCATGCCGCCCAATGTGCTTGCCCTCTTG	3292
Qy	2623	TTTCAAGCCTCAAGTGCCTCAATGGTGTGAGAGGCCCAGGACCACTGCAGATGCTCTCTC	2682
Db	3293	TTTCAAGCCTCAAGTGCCTCAATGGTGTGAGAGGCCCAGGACCACTGCAGATGCTCTCTC	3352
Qy	2683	TTTCAAGTCCATCAGCCCCATCAAGCCTTTTCTGCAGCCAGCAGTCTTGACAACCTATCTGG	2742
Db	3353	TTTCAAGTCCATCAGCCCCATCAAGCCTTTTCTGCAGCCAGCAGTCTTGACAACCTATCTGG	3412

Qy	2743	GAGTTTTTCAGAACTGCTCTTCAGTAGTTAGTTCAAGTGGACAGAGGGTGCTTCCAGTTT	2802
Db	3413	GAGTTTTTCAGAACTGCTCTTCAGTAGTTAGTTCAAGTGGACAGAGGGTGCTTCCAGTTT	3472
Qy	2803	GGACAAAAGGAGTTCAGGAGTAGTTTTAGGACATACCTCAATTGCTGAAGGAATCTTGG	2862
Db	3473	GGACAAAAGGAGTTCAGGAGTAGTTTTAGGACATACCTCAATTGCTGAAGGAATCTTGG	3532
Qy	2863	ACTTGAGCACCTAATGGATATATTTTCAGAGAGACAGATCACTTTGGATGATATTAGTTGA	2922
Db	3533	ACTTGAGCACCTAATGGATATATTTTCAGAGAGACAGATCACTTTGGATGATATTAGTTGA	3592
Qy	2923	GATGGGGCACAAGGAGCTGAAGAGATTGGAAATCAATGCTTTATGGACATAGGCACAACT	2982
Db	3593	GATGGGGCACAAGGAGCTGAAGAGATTGGAAATCAATGCTTTATGGACATAGGCACAACT	3652
Qy	2983	AAITTAAGGAGTCGAGAGACTTATCTCCGGACACAAGGTCCTTAACCCATATTAACTTT	3042
Db	3653	AAITTAAGGAGTCGAGAGACTTATCTCCGGACACAAGGTCCTTAACCCATATTAACTTT	3712
Qy	3043	GAACACCTCTGGTAGTGGAACTTCTTTATAGATCTGCTCCTCATGATATAAGAGTTTCA	3102
Db	3713	GAACACCTCTGGTAGTGGAACTTCTTTATAGATCTGCTCCTCATGATATAAGAGTTTCA	3772
Qy	3103	GTCGTGGAGGAAGAGATGCAAAAGTACAGTTCGAGAGACAGAGATGGAGTCAATGCAGG	3162
Db	3773	GTCGTGGAGGAAGAGATGCAAAAGTACAGTTCGAGAGACAGAGATGGAGTCAATGCAGG	3832
Qy	3163	TGGAATCTTCAACAGATACAAATTTCTCAAGATTTCAGAAAGTTGTATACAGAAACTATG	3222
Db	3833	TGGAATCTTCAACAGATACAAATTTCTCAAGATTTCAGAAAGTTGTATACAGAAACTATG	3892
Qy	3223	GGAAGATACACTCACCGGAGAAAAGTCTTCTGAAGAAAACACCAACCATGCGCAATGA	3282
Db	3893	GGAAGATACACTCACCGGAGAAAAGTCTTCTGAAGAAAACACCAACCATGCGCAATGA	3952
Qy	3283	ACGAATGCTATTTTCATGGGTCTCCTTTTCTGAATGCAATTTATCCAAAGGCTTTGATGA	3342
Db	3953	ACGAATGCTATTTTCATGGGTCTCCTTTTCTGAATGCAATTTATCCAAAGGCTTTGATGA	4012
Qy	3343	AAGCATCGGTACATAGGTGGTATGTTTCGAGCTGGCATTTATTTGCTGAAAACCTCTTC	3402
Db	4013	AAGCATCGGTACATAGGTGGTATGTTTCGAGCTGGCATTTATTTGCTGAAAACCTCTTC	4072
Qy	3403	CAAAAGCAATCAATATGATATGGAATTCGGAGAGGTACTGGGTGTCAGTTTCACAAAGA	3462
Db	4073	CAAAAGCAATCAATATGATATGGAATTCGGAGAGGTACTGGGTGTCAGTTTCACAAAGA	4132
Qy	3463	CAGATCTTGTATTTTCAGAGGAGAGCTGCTCTTTTCGCGGGTAACCTTTGGGAAAGTC	3522
Db	4133	CAGATCTTGTATTTTCAGAGGAGAGCTGCTCTTTTCGCGGGTAACCTTTGGGAAAGTC	4192
Qy	3523	TTTCTCGAGTTTCAGTGCATGAAATGGCAATTTCTCCTCCAGTGCATCACTCAGTCACT	3582
Db	4193	TTTCTCGAGTTTCAGTGCATGAAATGGCAATTTCTCCTCCAGTGCATCACTCAGTCACT	4252
Qy	3583	TGGTAGGCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACA	3642
Db	4253	TGGTAGGCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACA	4312
Qy	3643	GGCTTATCCTCAGTATTTAATTTACTTACAGATTTATGAGGCCCTGAAGGTATGGTCGATGG	3702
Db	4313	GGCTTATCCTCAGTATTTAATTTACTTACAGATTTATGAGGCCCTGAAGGTATGGTCGATGG	4372
Qy	3703	ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTGG	3762
Db	4373	ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAAATCATCAAGCAGCAGTGG	4432
Qy	3763	CCTCTACGTTTTTACTCCTTTGCTGAAAAAAA	3794
Db	4433	CCTCTACGTTTTTACTCCTTTGCTGAAAAAAA	4464

QY	643	TGGAGCTGAGCCAAACCATCGAAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC	702
Db	1323	TGGAGCTGAGCCAAACCATCGAAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC	1382
QY	703	ATCTGCCAAAGCAGTGTCTACTCGTGAATATAGAAGAAGTGAACCTCTTAGAAGATGCCAG	762
Db	1383	ATCTGCCAAAGCAGTGTCTACTCGTGAATATAGAAGAAGTGAACCTCTTAGAAGATGCCAG	1442
QY	763	GAGTGGCAATGAAGAAAATGATGCTCTACTCACACCATTAATATGTCACTGCCACGC	822
Db	1443	GAGTGGCAATGAAGAAAATGATGCTCTACTCACACCATTAATATGTCACTGCCACGC	1502
QY	823	AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATACACAGATAAAGAT	882
Db	1503	AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATACACAGATAAAGAT	1562
QY	883	TGTACAGCTGTACTGCAACATGGAGCTGATGTCCATGCTTAAGATAAAGTGATCTGGT	942
Db	1563	TGTACAGCTGTACTGCAACATGGAGCTGATGTCCATGCTTAAGATAAAGTGATCTGGT	1622
QY	943	ACCATTACAAATGCCGTGTCTTATGGTCATTATCAAGTAACTGAACCTTTTGGTCAAGCA	1002
Db	1623	ACCATTACAAATGCCGTGTCTTATGGTCATTATCAAGTAACTGAACCTTTTGGTCAAGCA	1062
QY	1003	TGGTGCTGTGTAATGCAATGAGCTGTGGCAATTCACCTCCTCTCATGAGCAGCTTC	1062
Db	1683	TGGTGCTGTGTAATGCAATGAGCTGTGGCAATTCACCTCCTCTCATGAGCAGCTTC	1122
QY	1063	TAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAAGTTATGTCGACAGCCACACCTGCT	1122
Db	1743	TAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAAGTTATGTCGACAGCCACACCTGCT	1802
QY	1123	CAATTGTACAATAAAGTGCTATAGACTTGGCTCCCAACCCACAGTTAAAGAAAGATT	1182
Db	1803	CAATTGTACAATAAAGTGCTATAGACTTGGCTCCCAACCCACAGTTAAAGAAAGATT	1862
QY	1183	AGCATATCAATTTAAAGGCCACTCGTTGCTGCAAGCTCCACGAAAGCTGATCTTACTCG	1242
Db	1863	AGCATATCAATTTAAAGGCCACTCGTTGCTGCAAGCTCCACGAAAGCTGATCTTACTCG	1922
QY	1243	AATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAACACATGAAC	1302
Db	1923	AATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAACACATGAAC	1982
QY	1303	AGCATTCATTTGCTGCTGCATCTCCATATCCCAAAAGAACAAATATGTGAACCTGT	1362
Db	1983	AGCATTCATTTGCTGCTGCATCTCCATATCCCAAAAGAACAAATATGTGAACCTGT	2042
QY	1363	GCTAAGAAAAGGAGCAACATCAATGAAAAGACTTAAAGATTTCTTGACTCCTCTGCACGT	1422
Db	2043	GCTAAGAAAAGGAGCAACATCAATGAAAAGACTTAAAGATTTCTTGACTCCTCTGCACGT	2102
QY	1423	GGCATCTGAGAAAGCTCATAAATGATCTGTTGAAGTACTGGTGAACATGAACCAAGGT	1482
Db	2103	GGCATCTGAGAAAGCTCATAAATGATCTGTTGAAGTACTGGTGAACATGAACCAAGGT	2162
QY	1483	TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTCACTCT	1542
Db	2163	TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTCACTCT	2222
QY	1543	ACAAACCTGCCGCTACTCTCTGAGCTATGGGTGTGATTCCTAACATTAATATCCCTTCAGGG	1602
Db	2223	ACAAACCTGCCGCTACTCTCTGAGCTATGGGTGTGATTCCTAACATTAATATCCCTTCAGGG	2282
QY	1603	CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCTCTCCACAGAGGATATCTC	1662
Db	2283	CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCTCTCCACAGAGGATATCTC	2342
QY	1663	ATTAGGTAAATCAGAGCGACAGACAAATGCTGGAAGCTGCCAAAGGCTGGAGATGTGGA	1722
Db	2343	ATTAGGTAAATCAGAGCGACAGACAAATGCTGGAAGCTGCCAAAGGCTGGAGATGTGGA	2402

QY	1723	AACTGTA	AAAAAACTGTGTTACTGTTCCAGAGTGTCAACTGCAGAGACATTTGAAGGCGGTCA	1781
DB	2403	AACTGT	AAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGGTCA	2462
QY	1783	GTCTAC	ACCACCTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCCT	1842
DB	2463	GTCTAC	ACCACCTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCCT	2522
QY	1843	ACAGCAT	GGAGCTGATGTGCATCTAAAGATAAAGAGAGGCGCTTGTACCTTTTGCACAATGC	1902
DB	2523	ACAGCAT	GGAGCTGATGTGCATCTAAAGATAAAGAGAGGCGCTTGTACCTTTTGCACAATGC	2582
QY	1903	ATGTCT	TATATGGACATTTAAAGTTTGCAGACTTCTTGTATAACATGGAGCAGTAGTTAA	1962
DB	2583	ATGTCT	TATATGGACATTTAAAGTTTGCAGACTTCTTGTATAACATGGAGCAGTAGTTAA	2642
QY	1963	TGTAGCT	GATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2022
DB	2643	TGTAGCT	GATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2702
QY	2023	AAITTTG	CAAACTTCTGCTCCAGCATGTGCGACCCCTACCAAAAAACACAGGGATGGAAA	2082
DB	2703	AAITTTG	CAAACTTCTGCTCCAGCATGTGCGACCCCTACCAAAAAACACAGGGATGGAAA	2762
QY	2083	TACTCCT	TTGGATCTGTTAAAGATGGATACAGATATCAAGATCTTGCTTAGGGGAGA	2142
DB	2763	TACTCCT	TTGGATCTGTTAAAGATGGATACAGATATCAAGATCTTGCTTAGGGGAGA	2822
QY	2143	TGCAGCT	TTTGCATAGTCTGCCAAGAGGGTTGTTTAGCCAGAGTGGAAGTTGTCTTC	2202
DB	2823	TGCAGCT	TTTGCATAGTCTGCCAAGAGGGTTGTTTAGCCAGAGTGGAAGTTGTCTTC	2882
QY	2203	TCCTGATA	ATGTAAATTTGCCGGCATACCCACAGCAGACATTCACACCTTTACATTTAGC	2262
DB	2883	TCCTGATA	ATGTAAATTTGCCGGCATACCCACAGCAGACATTCACACCTTTACATTTAGC	2942
QY	2263	AGCTGGT	TATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACCGAGCTGATGTGA	2322
DB	2943	AGCTGGT	TATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACCGAGCTGATGTGA	3002
QY	2323	TGCCCA	GACAAAAGGAGGACTTATTCCTTTACATAAATCGACATCTTTACGGGATGTAGA	2382
DB	3003	TGCCCA	GACAAAAGGAGGACTTATTCCTTTACATAAATCGACATCTTTACGGGATGTAGA	3062
QY	2383	TGTAGC	AGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTT	2442
DB	3063	TGTAGC	AGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTT	3122
QY	2443	CACACCT	TTTCCAGAGCAGCCCAAAAGGAGCAGACAGCTTTTGCTTTTGTTGCTAGC	2502
DB	3123	CACACCT	TTTCCAGAGCAGCCCAAAAGGAGCAGACAGCTTTTGCTTTTGTTGCTAGC	3182
QY	2503	CCATGG	AGCTTGACCCGACTCTTAAANAATCAGAGGACAAACACCTTTAGATTTAGTTTC	2562
DB	3183	CCATGG	AGCTTGACCCGACTCTTAAANAATCAGAGGACAAACACCTTTAGATTTAGTTTC	3242
QY	2563	ACCGAT	TGATGTACGGCTCTTCTGCAGACAGCATGCCGCCATCTGCTCTGCCCTCTTG	2622
DB	3243	ACCGAT	TGATGTACGGCTCTTCTGCAGACAGCATGCCGCCATCTGCTCTGCCCTCTTG	3302
QY	2623	TTACA	AGCCTCAAGTCTCAATGGTTGAGAGGCCACAGGAGCCACTGCAGATGCTCTCTC	2682
DB	3303	TTACA	AGCCTCAAGTCTCAATGGTTGAGAGGCCACAGGAGCCACTGCAGATGCTCTCTC	3362
QY	2683	TTCAGGT	TCCATCTAGCCCATCAAGCCTTTCTTGCAGCCAGCAGTCTTGACAACCTTATCTGG	2742
DB	3363	TTCAGGT	TCCATCTAGCCCATCAAGCCTTTCTTGCAGCCAGCAGTCTTGACAACCTTATCTGG	3422
QY	2743	GAGTTTT	TCAGACACTGTCTTTCAGTAGTTAGTTCAAGTTGGAAACAGAGGGTCTTCCAGTTT	2802
DB	3423	GAGTTTT	TCAGACACTGTCTTTCAGTAGTTAGTTCAAGTTGGAAACAGAGGGTCTTCCAGTTT	3482
QY	2803	GGAAAA	AGGAGGTTTCCAGGAGTAGATTTTATGCAATAACTCAATTCGTATAGGAATCTTGG	2862

Db 3483 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3542  
OY 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGTATTAGTTGA 2922  
Db 3543 ACTTGAGCACCTTAATGGATATATTTGAGAGAAACAGATCACTTTGGATGTATTAGTTGA 3602  
OY 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAAATCAATCTTATGGACATAGGCACAACT 2982  
Db 3603 GATGGGGCACAAGGAGCTGAAGGAGATTGGAAATCAATCTTATGGACATAGGCACAACT 3662  
OY 2983 AATTAAAGGAGTCGAGACATTAATCTCGGACAAACAGGCTTTAACCATAATTTAACTTT 3042  
Db 3663 AATTAAAGGAGTCGAGACATTAATCTCGGACAAACAGGCTTTAACCATAATTTAACTTT 3722  
OY 3043 GAACACCTCTGGTAGTGAACAATCTTATAGATCTGCTCTCATGATGAAGAGTTTCA 3102  
Db 3723 GAACACCTCTGGTAGTGAACAATCTTATAGATCTGCTCTCATGATGAAGAGTTTCA 3782  
OY 3103 GTCTGTGAGCAAGAGATGCAAACTACAGTCGAGACACAGAGATGGAGTTCATGCAGG 3162  
Db 3783 GTCTGTGAGCAAGAGATGCAAACTACAGTCGAGACACAGAGATGGAGTTCATGCAGG 3842  
OY 3163 TGGAACTCTTCAACAGATACATATCTCAAGATTCAAGAGTTTGTAAACAGAACTATG 3222  
Db 3843 TGGAACTCTTCAACAGATACATATCTCAAGATTCAAGAGTTTGTAAACAGAACTATG 3902  
OY 3223 GGAAGATACACTCACCGGAGAAAGATTTCTGAAGAAACACCAACCATGCCAATGA 3282  
Db 3903 GGAAGATACACTCACCGGAGAAAGATTTCTGAAGAAACACCAACCATGCCAATGA 3962  
OY 3283 ACGAATGCTATTTATGCGGCTCTCTTTGGAATGCAATATATCCAAAGGCTTTGATGA 3342  
Db 3963 ACGAATGCTATTTATGCGGCTCTCTTTGGAATGCAATATATCCAAAGGCTTTGATGA 4022  
OY 3343 AAGCATGCGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGGTGAACCTCTTC 3402  
Db 4023 AAGCATGCGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGGTGAACCTCTTC 4082  
OY 3403 CAAAGCAATCAATATGATGGAATTTGGAGAGTACTGGGTGTCAGTTTCACAAAGA 3462  
Db 4083 CAAAGCAATCAATATGATGGAATTTGGAGAGTACTGGGTGTCAGTTTCACAAAGA 4142  
OY 3463 CAGATCTTGTTACATTTGCCACAGGAGCTGCTCTTTTCCCGGTTAACTTGGGAAGTC 3522  
Db 4143 CAGATCTTGTTACATTTGCCACAGGAGCTGCTCTTTTCCCGGTTAACTTGGGAAGTC 4202  
OY 3523 TTTCTCTGAGTTCAGTGCATGAATGGAATGACATTTCTCTCCAGGTTCATCTCAGTCAC 3582  
Db 4203 TTTCTCTGAGTTCAGTGCATGAATGGAATGACATTTCTCTCCAGGTTCATCTCAGTCAC 4262  
OY 3583 TGGTAGGCCCCAGTGTAAATGGCTAGCATTAGCTGAATATGAGGCTGAAGGTATGGTCGATGG 3702  
Db 4263 TGGTAGGCCCCAGTGTAAATGGCTAGCATTAGCTGAATATGAGGCTGAAGGTATGGTCGATGG 4382  
OY 3703 ATAAATAGTATTTTAAAGAACTAATCCACTGAACCTTAAATCATCAAGACAGCAGTGG 3762  
Db 4383 ATAAATAGTATTTTAAAGAACTAATCCACTGAACCTTAAATCATCAAGACAGCAGTGG 4442  
OY 3763 CCTCTACGTTTTTACTCTCTTGTGTAAGAAAAA 3794  
Db 4443 CCTCTACGTTTTTACTCTCTTGTGTAAGAAAAA 4474

RESULT 8

LOCUS

DEFINITION

ACCESSION

AF264912 5810 bp mRNA linear PRI 14-MAY-2001  
Homo sapiens tankyrase-like protein (TNKL) mRNA, complete cds.  
AF264912

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## gene

## CDS

AF264912.1 GI:12005975

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 5810)

Kulimov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K.,

Scanlan, M.J., Jongeneel, C.V., Lagarkova, M.A., and Nedospasov, S.A.

Cloning and characterization of TNKL, a member of tankyrase gene

family

Genes Immun. 2 (1), 52-55 (2001)

21190090

11294570

2 (bases 1 to 5810)

Kulimov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K.,

Kashikarova, U.A., Boitchenko, V.E., Nedospasov, S.A. and

Lagarkova, M.A.

Direct Submission

Submitted (08-MAY-2000) Laboratory of Molecular Immunology, A.N.

Belozersky Institute of Physico-Chemical Biology, Moscow State

University, Moscow 119899, Russia

Location/Qualifiers

1..5810

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="10"

/map="10q22-10q24"

/tissue\_type="breast carcinoma"

/note="derived from clones MO-BC-203 and I203"

&lt;1..5810

/gene="TNKL"

/note="a new member of tankyrase gene family; current

UniGene cluster Hs.168491"

1..3501

/gene="TNKL"

/note="a new member of tankyrase family"

/codon\_start=1

/product="tankyrase-like protein"

/protein\_id="AAG44694.1"

/db\_xref="GI:12005976"

/translation="MSGRRCAAGGAGCAACAAAEPAARELFECAPNGDVERKRLV

TPKVNSTRDTAGKSTPLHFAAGFGKRDVVEYLLONGANVQARDDGGLPLHNACSF

HAEVNLLRHGADPNARDNNYTPLEAAIKKIDVICVILQHGAEPTIRNTDGRGA

LDLADPSKAVLTGEYKKDELLESARGNEEKMMALLPLNVNCHASDGRKSTPLHLA

AGYRVKIVQLLQHGADVHAKDKGDLVPLHNACSYGHVEYTELLVKGACVNMADLW

QLPLHEAANKRNEVCSLLLSYGAOPLLNCHNSAIDLAIPQLKRLAYEFKQHS

LLQAAEADVTRIKHLSEMVNFQHPQTHETALHCAASPYPKQKICELLRKGAN

INFKTEFTPLHVASEKAHNDVVEVVKHEAKVNDLNGOTSLHRAAYCGHQTGR

LLLSYGDPNIIISIQGTALQMGNEVQQLQSGISLGNSEADRLLEAKAGDVEV

KKLCTVQSNCRDIEGRQSTPLHFAAGYRVSVVEYLLQHGADVHAKDKGGLVPLHNA

CSYGHVEAELLVKGAVNVADLWKTPLHEAAKGYEICKLLQHGADPTKKNRD

GNTPDLVDKGDITDQLRGDAALDAAKKGCLARVKKLSSPDVNCRDTCGRHSTP

LHLAAGNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA

PDLWFTPLHEAAKGRQTQICALLAHGADPTLKNOEGOTPLDLSADVDVSAALLTAAM

PSPALPCYKQVINGVRSPGATADALSSGSPSSLSAASIDNLSGSFSELSSVVS

SSGTEGASLLEKEVPGVDFSTIQVFNGLGLHMDIFEREQITLDLVEMGHKELKE

IGINAYGHRHKLKIGVERLISQOGLNPILTLNTSGSTILIDLSPDDKEFQSVBEEM

QSTVREHGGHAGGIFNRYNLIKQKNKKLWERYTHRRKEVEENHNANERMLF

HGSFPVNAIHKGDFERHAYIGMGAGIYFAENSQKNOYVYGGGTGCGPVDKDRS

CICHKQLLFCRVTLGKSLQFSAMKMAHSPPHSHSVTGRPSVNGLALAEYVYIRGEQ

AYPELITYQIMRPEGVDC"

complement(1132..1277)

/note="SHGC-53837"

complement(3665..3975)

/note="WI-16054"

complement(5544..5777)

/note="WI-8339"

BASE COUNT 1736 a 1094 c 1231 g 1749 t

ORIGIN

Query Match 89.2%; Score 3387.2; DB 9; Length 5810;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	403	AGGTTTGGCGGAAAGACGTAGTCTGAATATTTGCTTCAGAAATGGTGCAGAGTGCCTCAAGC	462
Db	198	AGGTTTGGCGGAAAGACGTAGTCTGAATATTTGCTTCAGAAATGGTGCAGAGTGCCTCAAGC	257
Qy	463	ACGTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT	522
Db	258	ACGTGATGATGGGGGCTTATTCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT	317
Qy	523	AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAAATTTGGAATATATAC	582
Db	318	AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAAATTTGGAATATATAC	377
Qy	583	TCCTCTCCATGAAGTGCATTAAGAGAAAGATGATGTTTGGCAATTTGCTGTTACAGCA	642
Db	378	TCCTCTCCATGAAGTGCATTAAGAGAAAGATGATGTTTGGCAATTTGCTGTTACAGCA	437
Qy	643	TGGAGCTGAGCCAAACCATCCGAATATACAGATGAAGGACAGCATTTGGATTAGCAGATCC	702
Db	438	TGGAGCTGAGCCAAACCATCCGAATATACAGATGAAGGACAGCATTTGGATTAGCAGATCC	497
Qy	703	ATCTGCCAAACGCTGCTTACTGCTGGAATATAGAAAAGATGAACCTCTTAGAAAAGTCCAG	762
Db	498	ATCTGCCAAACGCTGCTTACTGCTGGAATATAGAAAAGATGAACCTCTTAGAAAAGTCCAG	557
Qy	763	GAGTGGCAATGAAGAAAATATGATGGCTCTACTCACACCAATTAATGTCACCTGCCACGC	822
Db	558	GAGTGGCAATGAAGAAAATATGATGGCTCTACTCACACCAATTAATGTCACCTGCCACGC	617
Qy	823	RAGTCATGCGAGAAGTCAACTCCATACATTTGGCAGCAGGATATAACAGAGTAAAGAT	882
Db	618	RAGTCATGCGAGAAGTCAACTCCATACATTTGGCAGCAGGATATAACAGAGTAAAGAT	677
Qy	883	TGTACAGCTGTTACTGCAACATGGAGCTGATGTCATGCTTAAGATTAAGGTGATCTGGT	942
Db	678	TGTACAGCTGTTACTGCAACATGGAGCTGATGTCATGCTTAAGATTAAGGTGATCTGGT	737
Qy	943	ACCATTTACACATGCTGCTTCTTATGTCATATATGAAGTAACTGAACCTTTTGGTCAAGCA	1002
Db	738	ACCATTTACACATGCTGCTTCTTATGTCATATATGAAGTAACTGAACCTTTTGGTCAAGCA	797
Qy	1003	TGCTGCTGTGTAATGCAATGGACTTTGGCAATTCACCTCTTCAATGAGGAGCTTC	1062
Db	798	TGCTGCTGTGTAATGCAATGGACTTTGGCAATTCACCTCTTCAATGAGGAGCTTC	857
Qy	1063	TAGAACAGGGTTGAAGTATGTTCTTCTTAAGTTATGTCGAGACCCCAACACTGCT	1122
Db	858	TAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGTCGAGACCCCAACACTGCT	917
Qy	1123	CAATTGTCACATAAAGTGCTATAGACTTGGCTCCCAACACACAGTTTAAAGAAAGATT	1182
Db	918	CAATTGTCACATAAAGTGCTATAGACTTGGCTCCCAACACACAGTTTAAAGAAAGATT	977
Qy	1183	AGCATATGAATTTAAGGGCACTCGTGTGCAAGTGCACAGAGCTGATGTTACTCG	1242
Db	978	AGCATATGAATTTAAGGGCACTCGTGTGCAAGTGCACAGAGCTGATGTTACTCG	1037
Qy	1243	AATCAAAAACATCTCTCTCGAAATGTTGAATTTCAAGCATCCTCAAAACACATGAAC	1302
Db	1038	AATCAAAAACATCTCTCTCGAAATGTTGAATTTCAAGCATCCTCAAAACACATGAAC	1097
Qy	1303	AGCATTCGATTTGCTGCTCATCTCCATATCCCAAGAAAGAAAGCAATATGTAAGT	1362
Db	1098	AGCATTCGATTTGCTGCTCATCTCCATATCCCAAGAAAGAAAGCAATATGTAAGT	1157
Qy	1363	GCTAAGAAAAGGAGCAACATCAATGAAGACTTAAGAAATTTTGACTCCTCTGCACGT	1422
Db	1158	GCTAAGAAAAGGAGCAACATCAATGAAGACTTAAGAAATTTTGACTCCTCTGCACGT	1217
Qy	1423	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT	1482
Db	1218	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT	1277











1004 GTGCGCTGTGTAATGCAATGGACCTGTGGCAATTCACCTCCCTTCATGAGGACGCTTCT 1063  
1061 GGTGCGTGTGTAATGCAATGGACCTGTGGCAATTCACCTCCCTTCATGAGGACGCTTCT 660  
1064 AAGAACAGGGTGAAGTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123  
661 AAGAACAGGGTGAAGTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
1124 AATTGTCACAATATAAGTGTCTATAGACTTGGCTCCACACACAGTAAAGAAAGATTA 1183  
721 AATTGTCACAATATAAGTGTCTATAGACTTGGCTCCACACACAGTAAAGAAAGATTA 780  
1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAGAGCTGATGTTACTCGA 1243  
781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAGAGCTGATGTTACTCGA 840  
1244 ATCAAAAACATCTCTCTCTGCAATGTTGAATTTCAAGCATCTCCCAACACATGAACA 1303  
841 ATCAAAAACATCTCTCTCTGCAATGTTGAATTTCAAGCATCTCCCAACACATGAACA 900  
1304 GCATTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1363  
901 GCATTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
1364 CTAAGAAAAGGAGCAACATCAATGAAAGACTTAAGAAATTTCTGACTCCTCTGCAAGTG 1423  
961 CTAAGAAAAGGAGCAACATCAATGAAAGACTTAAGAAATTTCTGACTCCTCTGCAAGTG 1020  
1424 GCATCTGAGAAAGCTCAATATGATGTTGTTGAAGTGTGTTGAACATCAAGCAAGGTT 1483  
1021 GCATCTGAGAAAGCTCAATATGATGTTGTTGAAGTGTGTTGAACATCAAGCAAGGTT 1080  
1484 AATGCTCTGGATAATCTTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCTA 1543  
1081 AATGCTCTGGATAATCTTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATCTA 1140  
1544 CAAACCTCGCCGCTACTCCTGAGCTATGGGTGTGATCTTAACATTTATATCCCTTCAGGC 1603  
1141 CAAACCTCGCCGCTACTCCTGAGCTATGGGTGTGATCTTAACATTTATATCCCTTCAGGC 1200  
1604 TTTACTGCTTTACAGATGGGAATGAAATGACAGCACTCTCCACAGAGGTTATCTCA 1663  
1201 TTTACTGCTTTACAGATGGGAATGAAATGACAGCACTCTCCACAGAGGTTATCTCA 1260  
1664 TTAGTGAATTCAGAGGCAAGACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAA 1723  
1261 TTAGTGAATTCAGAGGCAAGACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAA 1320  
1724 ACTGTAAAAAACTGTGTACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGGCTCAG 1783  
1321 ACTGTAAAAAACTGTGTACTGTTGAGAGTGTCAACTGCAGAGACATTTGAAGGGGCTCAG 1380  
1784 TCTACACCACTTCATTTGCAAGTGGGTATACAGAGTGTCCGTTGGTGAATATCTGCTA 1843  
1381 TCTACACCACTTCATTTGCAAGTGGGTATACAGAGTGTCCGTTGGTGAATATCTGCTA 1440  
1844 CAGCATGAGCTGTGATGCTGCTAAAGATAAAGAGAGGCTTTGTACCTTTGCACAAATGCA 1903  
1441 CAGCATGAGCTGTGATGCTGCTAAAGATAAAGAGAGGCTTTGTACCTTTGCACAAATGCA 1500  
1904 TGTCTTATGGACATATGCAAGTTCGCAACTTCTGTTAAACATGGAGCAGTATGTTAAT 1963  
1501 TGTCTTATGGACATATGCAAGTTCGCAACTTCTGTTAAACATGGAGCAGTATGTTAAT 1560  
1964 GTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAA 2023  
1561 GTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAA 1620  
2024 ATTTGCAAACTTCTCTCCAGCATGGTGCAGACCTTACCAAAAAACAGGAGTGAAT 2083  
1621 ATTTGCAAACTTCTCTCCAGCATGGTGCAGACCTTACCAAAAAACAGGAGTGAAT 1680  
2084 ACTCCTTTGGATCTTTGTAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGNT 2143

1681 ACTCCTTTGGATCTTTGTTAAAGATGAGATACAGATATTCATTTATCTGCTTAGGGAGAT 1740  
2144 GCAGCTTTGCTAGATGCTGCCAAGAGAGGCTTCTTAGCCAGAGTGAAGAGTGTCTTCT 2203  
1741 GCAGCTTTGCTAGATGCTGCCAAGAGAGGCTTCTTAGCCAGAGTGAAGAGTGTCTTCT 1800  
2204 CCTGATATGTAATTTGCCCGCATCCCAAGGAGACACATTTCAACACCTTTACATTTAGCA 2263  
1801 CCTGATATGTAATTTGCCCGCATCCCAAGGAGACACATTTCAACACCTTTACATTTAGCA 1860  
2264 GCTGCTTATATAATTTAGAACTTGCAGAGTATTTGTTTACAACAGGAGCTGATGTGAAT 2323  
1861 GCTGCTTATATAATTTAGAACTTGCAGAGTATTTGTTTACAACAGGAGCTGATGTGAAT 1920  
2324 GCCCAGACAAAGGAGGACTTATCTTTACATAATGACAGCATCTTACGGGATGTAGAT 2383  
1921 GCCCAGACAAAGGAGGACTTATCTTTACATAATGACAGCATCTTACGGGATGTAGAT 1980  
2384 GTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCACGACAAATGGGCTTTC 2443  
1981 GTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCACGACAAATGGGCTTTC 2040  
2444 ACACCTTTGCACGAGCAGGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGTGCTAGCC 2503  
2041 ACACCTTTGCACGAGCAGGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGTGCTAGCC 2100  
2504 CATGAGCTGACCCGACTCTTAAATAATCAGGAAGACAAACACCTTTAGATTTAGTTTCA 2563  
2101 CATGAGCTGACCCGACTCTTAAATAATCAGGAAGACAAACACCTTTAGATTTAGTTTCA 2160  
2564 GCGGATGATGTACAGGCTCTCTTGACAGCAGCCATGCCCCCATCTGCTGCGCCCTTCT 2623  
2161 GCGGATGATGTACAGGCTCTCTTGACAGCAGCCATGCCCCCATCTGCTGCGCCCTTCT 2220  
2624 TACAGCCTCAAGTCTCAATGGTGTGAGAACCCAGGAGCCACTGCAGATGCTCTCTCT 2683  
2221 TACAGCCTCAAGTCTCAATGGTGTGAGAACCCAGGAGCCACTGCAGATGCTCTCTCT 2280  
2684 TCAGTCTCATCTAGCCCATCAAGCCTTTCTGACAGCAGCCACTGCACAACTTATCTGGG 2743  
2281 TCAGTCTCATCTAGCCCATCAAGCCTTTCTGACAGCAGCCACTGCACAACTTATCTGGG 2340  
2744 AGTTTTTCAAGACTGCTCTAGTATTTAGTCAAGTGGAAACAGAGGCTTCCAGTTTG 2803  
2341 AGTTTTTCAAGACTGCTCTAGTATTTAGTCAAGTGGAAACAGAGGCTTCCAGTTTG 2400  
2804 GAGAAAAGGAGTTCCAGGAGTATTTAGCATAACTCAATTCGTAAAGGAATCTTGA 2863  
2401 GAGAAAAGGAGTTCCAGGAGTATTTAGCATAACTCAATTCGTAAAGGAATCTTGA 2460  
2864 CTTGAGCACCTAATCGATATATTTGAGAGAACAGATCACTTTGGATGATTTAGTTGAG 2923  
2461 CTTGAGCACCTAATCGATATATTTGAGAGAACAGATCACTTTGGATGATTTAGTTGAG 2520  
2924 ATGGGCAAGAGGAGCTGAAGGAGATTTGAATCAATGCTTTATGGACATAGGCAAACTA 2983  
2521 ATGGGCAAGAGGAGCTGAAGGAGATTTGAATCAATGCTTTATGGACATAGGCAAACTA 2580  
2984 ATTAAGGAGTGGAGAGCTTATCTCCGACAAAGGCTTTAACCCATATTTAACTTTG 3043  
2581 ATTAAGGAGTGGAGAGCTTATCTCCGACAAAGGCTTTAACCCATATTTAACTTTG 2640  
3044 AACACCTCTGTTAGTGGAACTTCTTTATAGATCTCTCTCTCTGATGATAAGAGTTTCAG 3103  
2641 AACACCTCTGTTAGTGGAACTTCTTTATAGATCTCTCTCTGATGATAAGAGTTTCAG 2700  
3104 TCTGTGGAGGAAGATGCAAGTACAGTTCAGAGCAGACAGAGTGGAGTGCATGCAGT 3163  
2701 TCTGTGGAGGAAGATGCAAGTACAGTTCAGAGCAGACAGAGTGGAGTGCATGCAGT 2760  
3164 GGAATCTTCAACAGATACAAATTTCTCAAGATTCAGAGGTTTGTAAACAAAGAACTATGG 3223

Db 2761 GAACTCTCAACAGATACAAATATTTCTCAAGATTTCAGAAAGTTTGTAAACAAGAACTATGG 2820  
Qy 3224 GAAAGATACACTCACCAGGAGAAAAGAGTTTCTGAAGAAAACACACCAATGCCAATGAA 3283  
Db 2821 GAAAGATACACTCACCAGGAGAAAAGAGTTTCTGAAGAAAACACACCAATGCCAATGAA 2880  
Qy 3284 GAAAGATACACTCACCAGGAGAAAAGAGTTTCTGAAGAAAACACACCAATGCCAATGAA 3343  
Db 2881 GAAAGATACACTCACCAGGAGAAAAGAGTTTCTGAAGAAAACACACCAATGCCAATGAA 2940  
Qy 3344 AGGCATCGGTACATAGAGTGGTATGTTGGAGCTGGCAATTTATTTTGGTGAAGAACTCTTCC 3403  
Db 2941 AGGCATCGGTACATAGAGTGGTATGTTGGAGCTGGCAATTTATTTTGGTGAAGAACTCTTCC 3000  
Qy 3404 AAAAGCAATCAATATGTAATGGAATTTGGAGAGGTACTGGTCTCCAGTTCACAAAGAC 3463  
Db 3001 AAAAGCAATCAATATGTAATGGAATTTGGAGAGGTACTGGTCTCCAGTTCACAAAGAC 3060  
Qy 3464 AGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCT 3523  
Db 3061 AGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCGGGTAACCTTTGGGAAAGTCT 3120  
Qy 3524 TTCTGCGAGTTCCAGTGAATGAAATGGCAGCATTTCTTCCAGGTCACTCAGTCACT 3583  
Db 3121 TTCTGCGAGTTCCAGTGAATGAAATGGCAGCATTTCTTCCAGGTCACTCAGTCACT 3180  
Qy 3584 GGTAGCCAGTGAATGGCTAGCATTTAGTGAATTTATTTTACAGAGGAGACAG 3643  
Db 3181 GGTAGCCAGTGAATGGCTAGCATTTAGTGAATTTATTTTACAGAGGAGACAG 3240  
Qy 3644 GCTTATCTCTGAGTATTTAATTAATTTACCAGATTTATGAGCCCTGAAGGTATGTCGATGGA 3703  
Db 3241 GCTTATCTCTGAGTATTTAATTAATTTACCAGATTTATGAGCCCTGAAGGTATGTCGATGGA 3300  
Qy 3704 TAAATAGTATTTTAAAGAACTAATTCACAGTGAACCTTAAATCATCAAGCAGAGTGGC 3763  
Db 3301 TAAATAGTATTTTAAAGAACTAATTCACAGTGAACCTTAAATCATCAAGCAGAGTGGC 3360  
Qy 3764 CTCTAGTTTACTCTTCTTGGTGAAGAAAAGAAA 3797  
Db 3361 CTCTAGTTTACTCTTCTTGGTGAAGAAAAGAAA 3394

RESULT 11  
AX076213  
LOCUS AX076213 4512 bp DNA linear PAT 06-FEB-2001  
DEFINITION Sequence 3 from Patent WO0104326.  
ACCESSION AX076213  
VERSION AX076213.1 GI:12710838  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4512)  
AUTHORS Berthelsen, J., Toma, S. and Isacchi, A.  
TITLE Tankyrase homolog protein (thp), nucleic acids, and methods related to the same  
JOURNAL Patent: WO 0104326-A 3 18-JAN-2001;  
PHARMACIA & UPJOHN S.p.A. (IT)  
FEATURES Location/Qualifiers  
source 1..4512  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1124  
/note="n is any base"  
2672  
/note="n is any base"  
BASE COUNT 1356 a 878 c 1002 g 1268 t 8 others  
ORIGIN

Query Match 89.1%; Score 3382.8; DB 6; Length 4512;  
Best Local Similarity 99.7%; Pred. No. 0;

Matches 3381; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
Qy 403 AGGTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAGATGGTGCAGTGTCCCAAGC 462  
Db 220 AGGTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAGATGGTGCAGTGTCCCAAGC 279  
Qy 463 ACCTGATGATGGGGGCTTATTTCTCTTTCATATGATGCTCTTTTGGTCACTGCTGAAGT 522  
Db 280 ACCTGATGATGGGGGCTTATTTCTCTTTCATATGATGCTCTTTTGGTCACTGCTGAAGT 339  
Qy 523 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATAC 582  
Db 340 AGTCAATCTCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATAC 399  
Qy 583 TCCTCTCATGAGCTGCAATTTAAAGGAAAGATTGATGTTTGGTCACTGCTGTGTACAGCA 642  
Db 400 TCCTCTCATGAGCTGCAATTTAAAGGAAAGATTGATGTTTGGTCACTGCTGTGTACAGCA 459  
Qy 643 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCAATTTGATTTAGCAGATCC 702  
Db 460 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCAATTTGATTTAGCAGATCC 519  
Qy 703 ATCTGCCAAAGCAGTGTCTTACTGCTGAATATAGAAAGATGAACTCTTAGAAGATGCCAG 762  
Db 520 ATCTGCCAAAGCAGTGTCTTACTGCTGAATATAGAAAGATGAACTCTTAGAAGATGCCAG 579  
Qy 763 GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 822  
Db 580 GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 639  
Qy 823 AAGTGTGGCAGAAAGTCAACTCCATTTACATTTTGGCAGCAGGATATTAACAGATAAGAT 882  
Db 640 AAGTGTGGCAGAAAGTCAACTCCATTTACATTTTGGCAGCAGGATATTAACAGATAAGAT 699  
Qy 883 TGTACACCTGTTACTGCAACATGGAGCTGATGCTCATGCTTAAAGATAAGGTGATCTGGT 942  
Db 700 TGTACACCTGTTACTGCAACATGGAGCTGATGCTCATGCTTAAAGATAAGGTGATCTGGT 759  
Qy 943 ACCATTACACAAATGCCTGTTCTTATGCTCATTTATGAAGTAACCTTTTGGTTCAGCA 1002  
Db 760 ACCATTACACAAATGCCTGTTCTTATGCTCATTTATGAAGTAACCTTTTGGTTCAGCA 819  
Qy 1003 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTTCATGAGGAGCTTC 1062  
Db 820 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTTCATGAGGAGCTTC 879  
Qy 1063 TAAGAACAGGCTTGAAGTATGCTCTCTTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 1122  
Db 880 TAAGAACAGGCTTGAAGTATGCTCTCTTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 939  
Qy 1123 CAATTGTCAACAATAAAAGTGTATAGACTTGGCTCCCAACACACAGTTAAAGAAAGATT 1182  
Db 940 CAATTGTCAACAATAAAAGTGTATAGACTTGGCTCCCAACACACAGTTAAAGAAAGATT 999  
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAAGCTGATTTACTCG 1242  
Db 1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAGAAGCTGATTTACTCG 1059  
Qy 1243 AATCAAAAACATCTCTCTCTGGAATGGTGAATTTTCAAGCATCTTCAACACATGAAC 1302  
Db 1060 AATCAAAAACATCTCTCTCTGGAATGGTGAATTTTCAAGCATCTTCAACACATGAAC 1119  
Qy 1303 AGCATTTGATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTTGAACTGTT 1362  
Db 1120 AGCATTTGATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTTGAACTGTT 1179  
Qy 1363 GCTAAGAAAAGCAGCAACATCAATGAAAAGACTTAAAGAAATTTCTTGACTCTCTGCACGT 1422  
Db 1180 GCTAAGAAAAGCAGCAACATCAATGAAAAGACTTAAAGAAATTTCTTGACTCTCTGCACGT 1239  
Qy 1423 GGCATCTCAGAAAAGCTCATATGATGTTTGAAGTAGTGTGAACATGAAGCAAAAGGT 1482  
Db 1240 GGCATCTCAGAAAAGCTCATATGATGTTTGAAGTAGTGTGAACATGAAGCAAAAGGT 1299

QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACAGAGCTGCAATATTTGTTGCTCATCT 1542  
DB 1300 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACAGAGCTGCAATATTTGTTGCTCATCT 1359  
QY 1543 ACAACCTGCGCCCTACTCTGAGCTATGCGTGTGATCCCTAAACATATATATCCCTTCAGGG 1602  
DB 1360 ACAACCTGCGCCCTACTCTGAGCTATGCGTGTGATCCCTAAACATATATATCCCTTCAGGG 1419  
QY 1603 CTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAAGAGGATCTTC 1662  
DB 1420 CTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAAGAGGATCTTC 1479  
QY 1663 ATTAGGTAATTCAGAGCGACAGACAATTTGCTGGAAGCTGCAAGAGCTGAGATGTGCA 1722  
DB 1480 ATTAGGTAATTCAGAGCGACAGACAATTTGCTGGAAGCTGCAAGAGCTGAGATGTGCA 1539  
QY 1723 AACTGTAATAAANAATGTTACTGTTACAGAGTGTCAACTGTCAGAGACATTTGAAGGGGCTCA 1782  
DB 1540 AACTGTAATAAANAATGTTACTGTTACAGAGTGTCAACTGTCAGAGACATTTGAAGGGGCTCA 1599  
QY 1783 GTCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGTTGGAATATCTGCT 1842  
DB 1600 GTCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGTTGGAATATCTGCT 1659  
QY 1843 ACAGCATGGAGCTGATGTGATGCTTAAAGATAAAGRRGSCCTTGTACCTTTGCAACATGC 1902  
DB 1660 ACAGCATGGAGCTGATGTGATGCTTAAAGATAAAGRRGSCCTTGTACCTTTGCAACATGC 1719  
QY 1903 ATGTTCTTATGGACATTTCAAGCTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTAA 1962  
DB 1720 ATGTTCTTATGGACATTTCAAGCTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTAA 1779  
QY 1963 TGTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATCA 2022  
DB 1780 TGTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATCA 1839  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGGAGTGGAAA 2082  
DB 1840 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGGAGTGGAAA 1899  
QY 2083 TACTCCTTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTCTTAGGGGAGA 2142  
DB 1900 TACTCCTTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTCTTAGGGGAGA 1959  
QY 2143 TGCAGCTTTGCTAGAGTGTCCCAAGAGGTTGTTTACAGAGTGAAGAGTGTCTTTC 2202  
DB 1960 TGCAGCTTTGCTAGAGTGTCCCAAGAGGTTGTTTACAGAGTGAAGAGTGTCTTTC 2019  
QY 2203 TCTGATAATGTAATTTGCGCGATACCAAGCAGACATTTCAACACCTTTTACATTTAGC 2262  
DB 2020 TCTGATAATGTAATTTGCGCGATACCAAGCAGACATTTCAACACCTTTTACATTTAGC 2079  
QY 2263 AGCTGGTTATTAATTTAGAAGTTCAGAGTATTTGTTTACACACGAGCTGATGTGAA 2322  
DB 2080 AGCTGGTTATTAATTTAGAAGTTCAGAGTATTTGTTTACACACGAGCTGATGTGAA 2139  
QY 2323 TGCCCAAGACAAGAGGAGCTTATTCCTTTTACATATATGACAGATCTTACGGGCATGTAGA 2382  
DB 2140 TGCCCAAGACAAGAGGAGCTTATTCCTTTTACATATATGACAGATCTTACGGGCATGTAGA 2199  
QY 2383 TGTAGCAGCTTACTAATAAAGTATTAATGATGTGTCAATGCGCAGCAAAATGGGCTTT 2442  
DB 2200 TGTAGCAGCTTACTAATAAAGTATTAATGATGTGTCAATGCGCAGCAAAATGGGCTTT 2259  
QY 2443 CACACCTTTTGCAGAGCAGCCCAAGGAGCAACAGAGCTTGTGCTTTTCTAGC 2502  
DB 2260 CACACCTTTTGCAGAGCAGCCCAAGGAGCAACAGAGCTTGTGCTTTTCTAGC 2319  
QY 2503 CCATGAGCTGACCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
DB 2320 CCATGAGCTGACCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2379

QY 2563 AGCGATGATGTACAGGCTCTTCTGACAGCAGCATGCCCATCTGCTGCTGCCCTCTG 2622  
DB 2380 AGCGATGATGTACAGGCTCTTCTGACAGCAGCATGCCCATCTGCTGCTGCCCTCTG 2439  
QY 2623 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGAGCCACTGAGATGCTCTCTC 2682  
DB 2440 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGAGCCACTGAGATGCTCTCTC 2499  
QY 2683 TTCAGTCCACTTACGCCATCAAGCCTTCTGACGCCAGAGCTCTTGACAATTAATCTTCG 2742  
DB 2500 TTCAGTCCACTTACGCCATCAAGCCTTCTGACGCCAGAGCTCTTGACAATTAATCTTCG 2559  
QY 2743 GAGTTTTTTCAGAACTGTCTTCAAGTAGTTAGTTCAAGTGTGAAGAGGCTCTCCAGTTT 2802  
DB 2560 GAGTTTTTTCAGAACTGTCTTCAAGTAGTTAGTTCAAGTGTGAAGAGGCTCTCCAGTTT 2619  
QY 2803 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGNAATCTTCG 2862  
DB 2620 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGNAATCTTCG 2679  
QY 2863 ACTTACAGCAGCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGTATTAGTTGA 2922  
DB 2680 ACTTACAGCAGCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGTATTAGTTGA 2739  
QY 2923 GATGGGACACAGGAGCTCAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2982  
DB 2740 GATGGGACACAGGAGCTCAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2799  
QY 2983 AATTAAGAGTTCGAGAGCTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTT 3042  
DB 2800 AATTAAGAGTTCGAGAGCTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTT 2859  
QY 3043 GAACACCTCTGCTAGTGAACAATTTCTATAGATCTGCTCTCTGATATAAGAGATTTC 3102  
DB 2860 GAACACCTCTGCTAGTGAACAATTTCTATAGATCTGCTCTCTGATATAAGAGATTTC 2919  
QY 3103 GTCTGTGGAGAGAGATGCAAGGTACAGTTTCGAGAGCACAGAGATGGAGTCAATGCAGG 3162  
DB 2920 GTCTGTGGAGAGAGATGCAAGGTACAGTTTCGAGAGCACAGAGATGGAGTCAATGCAGG 2979  
QY 3163 TGGATCTTCAACAGATACATATTTCTCAAGATTCTCAAGATTCTCAAGATTCTCAAGATTCT 3222  
DB 2980 TGGATCTTCAACAGATACATATTTCTCAAGATTCTCAAGATTCTCAAGATTCTCAAGATTCT 3039  
QY 3223 GGAAGATACACTCACCGGAGAAAGAGTTCTGAGAGAACACACCATGCAATGA 3282  
DB 3040 GGAAGATACACTCACCGGAGAAAGAGTTCTGAGAGAACACACCATGCAATGA 3099  
QY 3283 ACAGATGCTATTTTCATGGTCTCTCTTTTGTGAATGCAATTTATCCACAAAGGCTTTGATGA 3342  
DB 3100 ACAGATGCTATTTTCATGGTCTCTCTTTTGTGAATGCAATTTATCCACAAAGGCTTTGATGA 3159  
QY 3343 AAGCATGCTGATAGTGTATGTTTGGAGCTGGCATTTATTTGCTGAAACCTTCTC 3402  
DB 3160 AAGCATGCTGATAGTGTATGTTTGGAGCTGGCATTTATTTGCTGAAACCTTCTC 3219  
QY 3403 CAAAGCAATCAATATGATGGAATTTGAGGAGTACTGGGTGCTCCAGTTCACAAAGA 3462  
DB 3220 CAAAGCAATCAATATGATGGAATTTGAGGAGTACTGGGTGCTCCAGTTCACAAAGA 3279  
QY 3463 CAGATCTTGTACATTTGCCACAGGCTGCTCTTTTGGCGGGTAACCTTTGGGAAGTC 3522  
DB 3280 CAGATCTTGTACATTTGCCACAGGCTGCTCTTTTGGCGGGTAACCTTTGGGAAGTC 3339  
QY 3523 TTTCTCAGCTTACATTTGCAATGGAATTTAGCTGATATGTTTATTTACAGAGGAGAA 3582  
DB 3340 TTTCTCAGCTTACATTTGCAATGGAATTTAGCTGATATGTTTATTTACAGAGGAGAA 3399  
QY 3583 TGTAGGCTGCTGATTTAAATGGCTTAGCTGATATGTTTATTTACAGAGGAGAA 3642  
DB 3400 TGTAGGCTGCTGATTTAAATGGCTTAGCTGATATGTTTATTTACAGAGGAGAA 3459  
QY 3643 GGCTTATCTTGAGTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3702



QY 1482 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTTCATC 1541  
Db 1576 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTTCATC 1635  
QY 1542 TACAAACCTTGGCCCTACTCTGAGCTATGGGTGTGATCCCTAACATTTATATCCCTTCAGG 1601  
Db 1636 TACAAACCTTGGCCCTACTCTGAGCTATGGGTGTGATCCCTAACATTTATATCCCTTCAGG 1695  
QY 1602 GCTTTACTGCTTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCT 1661  
Db 1696 GCTTTACTGCTTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCT 1755  
QY 1662 CATTTAGTAAATTCAGAGGACAGACAAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCG 1721  
Db 1756 CATTTAGTAAATTCAGAGGACAGACAAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCG 1815  
QY 1722 AAATGTAAAAAATCTGTACTGTTTCAGAGTGTCAACTGCGACAGACATTTGAAGGGCGTC 1781  
Db 1816 AAATGTAAAAAATCTGTACTGTTTCAGAGTGTCAACTGCGACAGACATTTGAAGGGCGTC 1875  
QY 1782 AGTCTACACCACCTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGC 1841  
Db 1876 AGTCTACACCACCTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGC 1935  
QY 1842 TACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGAGGCCCTTGACCTTTGCACAATG 1901  
Db 1936 TACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGAGGCCCTTGACCTTTGCACAATG 1995  
QY 1902 CATGTTCTTATGACATATTAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTA 1961  
Db 1996 CATGTTCTTATGACATATTAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTA 2055  
QY 1962 ATCTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATG 2021  
Db 2056 ATCTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATG 2115  
QY 2022 AAATTTGCAAACTTCTGCTCCAGATGGTGCAGACCTTACCAAAAAACAGGAGTGGAA 2081  
Db 2116 AAATTTGCAAACTTCTGCTCCAGATGGTGCAGACCTTACCAAAAAACAGGAGTGGAA 2175  
QY 2082 ATACTCCTTTTGGATCTTGTAAAGATGGAGATACAGATATCAAGATCTCTTAGGGGAG 2141  
Db 2176 ATACTCCTTTTGGATCTTGTAAAGATGGAGATACAGATATCAAGATCTCTTAGGGGAG 2235  
QY 2142 ATGACAGCTTTTGTAGATGCTGCCAAGAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2201  
Db 2236 ATGACAGCTTTTGTAGATGCTGCCAAGAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2295  
QY 2202 CTCTGATAATGTAAATTTGCCGGATACCCAAGCAGACATTTCAACACCTTTACATTTAG 2261  
Db 2296 CTCTGATAATGTAAATTTGCCGGATACCCAAGCAGACATTTCAACACCTTTACATTTAG 2355  
QY 2262 CAGCTGTTATATAATTTTAAAGTTTGCAGAGTATTTGTTTACAAACAGGAGCTGATGTGA 2321  
Db 2356 CAGCTGTTATATAATTTTAAAGTTTGCAGAGTATTTGTTTACAAACAGGAGCTGATGTGA 2415  
QY 2322 ATGCCAAGCAAAAGGAGGACTTTATTCCTTTTACATATATGAGCAGCTTTACGGCATGTAG 2381  
Db 2416 ATGCCAAGCAAAAGGAGGACTTTATTCCTTTTACATATATGAGCAGCTTTACGGCATGTAG 2475  
QY 2382 ATGTAGCAGCTCTACTTAATAAGTATAATGATGTGCAATGCCAGGCAAAATGGGCTT 2441  
Db 2476 ATGTAGCAGCTCTACTTAATAAGTATAATGATGTGCAATGCCAGGCAAAATGGGCTT 2535  
QY 2442 TCACACCTTTTGCAGAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTGTAG 2501  
Db 2536 TCACACCTTTTGCAGAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTGTAG 2595  
QY 2502 CCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTTATGATTTAGTTT 2561  
Db 2596 CCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTTATGATTTAGTTT 2655  
QY 2562 CAGCGGATGATGACGGCTCTCTTGACAGCAGCCATGCCCCCATCTGCTCGCCCTCTT 2621

Db 2656 CAGCGGATGATGTCAGGGCTCTTCTGACAGCAGCCATGCCCATCTGCTCTGCCCTCTT 2715  
QY 2622 GTTTACAGGCTCTAAGTGCTCAATGGTGTGAGAAGCCACAGAGCCACTGCGAGTGCTCTCT 2681  
Db 2716 GTTTACAGGCTCTAAGTGCTCAATGGTGTGAGAAGCCACAGAGCCACTGCGAGTGCTCTCT 2775  
QY 2682 CTTTCAAGTCCATCTAGCCCATCAAGCCTTTTCTGACGCCAGCAGCTTTGACAACTTATCTG 2741  
Db 2776 CTTTCAAGTCCATCTAGCCCATCAAGCCTTTTCTGACGCCAGCAGCTTTGACAACTTATCTG 2835  
QY 2742 GGAGTTTTTTCAGAACTGTCTTCAAGTGTAGTTCAAGTGGAAACAGAGGGTGTCTCCAGTT 2801  
Db 2836 GGAGTTTTTTCAGAACTGTCTTCAAGTGTAGTTCAAGTGGAAACAGAGGGTGTCTCCAGTT 2895  
QY 2802 TGGAGAAAAAGAGGTTCCAGGAGTAGATTTTATAGCTAACTCAATTCGTAAGAAATCTTG 2861  
Db 2896 TGGAGAAAAAGAGGTTCCAGGAGTAGATTTTATAGCTAACTCAATTCGTAAGAAATCTTG 2955  
QY 2862 GACTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATAGTTG 2921  
Db 2956 GACTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATAGTTG 3015  
QY 2922 AGATGGGACACAGAGCTGAAAGGATTTGGAATCAATGCTTATGGACATAGGCACAAAC 2981  
Db 3016 AGATGGGACACAGAGCTGAAAGGATTTGGAATCAATGCTTATGGACATAGGCACAAAC 3075  
QY 2982 TAATTTAAAGGAGTCGAGAGACTTATCTCCGACAAACAGGCTTTAAOCCATATTTAACTT 3041  
Db 3076 TAATTTAAAGGAGTCGAGAGACTTATCTCCGACAAACAGGCTTTAAOCCATATTTAACTT 3135  
QY 3042 TGAACACCTCTGCTAGTGGAAACAATTTCTTATAGATCTGCTCTCTGATGATAAAGAGTTT 3101  
Db 3136 TGAACACCTCTGCTAGTGGAAACAATTTCTTATAGATCTGCTCTCTGATGATAAAGAGTTT 3195  
QY 3102 AGTCTGTGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGACAGATGGAGTTCATGCAG 3161  
Db 3196 AGTCTGTGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGACAGATGGAGTTCATGCAG 3255  
QY 3162 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTTCAAGAGTTTGTAAACAAGAAACTAT 3221  
Db 3256 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTTCAAGAGTTTGTAAACAAGAACTAT 3315  
QY 3222 GGAAAGATACATCTACCCGAGAGAAAGAGTTTCTGAAGAAACCAACCATGCCAATG 3281  
Db 3316 GGAAAGATACATCTACCCGAGAGAAAGAGTTTCTGAAGAAACCAACCATGCCAATG 3375  
QY 3282 AACGAATGCTATTTTCATGGGTCTCTTTTCTGGAATGCAATTTCCACAAAGGCTTTGATG 3341  
Db 3376 AACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATTTCCACAAAGGCTTTGATG 3435  
QY 3342 AAAGGATGCTACATAGGTGGTATGTTTGGAGTGGCATTTATTTTGTGAAACACTCTT 3401  
Db 3436 AAAGGATGCTACATAGGTGGTATGTTTGGAGTGGCATTTATTTTGTGCTGAAACACTCTT 3495  
QY 3402 CCAAAAGCAATCAATATGTATATGGAATTTGAGAGAGTACTGGGTGTCAGTTCACAAAG 3461  
Db 3496 CCAAAAGCAATCAATATGTATATGGAATTTGAGAGAGTACTGGGTGTCAGTTCACAAAG 3555  
QY 3462 ACAGATCTTCTTACATTTGCCACAGGAGTGTCTTTTTCGGGGTAACTTTGGGAAAGT 3521  
Db 3556 ACAGATCTTCTTACATTTGCCACAGGAGTGTCTTTTTCGGGGTAACTTTGGGAAAGT 3615  
QY 3522 CTTTCTCTGCAAGTTCAGTGCATGAAATGGCAATTTCTCTCCAGGTCATCAGTCA 3581  
Db 3616 CTTTCTCTGCAAGTTCAGTGCATGAAATGGCAATTTCTCTCCAGGTCATCAGTCA 3675  
QY 3582 CTGGTAGCCCGCAGTGAATGGCCCTAGCATTTAGCTGAATGTATTTTACAGAGGAGAAC 3641  
Db 3676 CTGGTAGCCCGCAGTGAATGGCCCTAGCATTTAGCTGAATGTATTTTACAGAGGAGAAC 3735  
QY 3642 AGGCTTATCTCTGAGTATTTAACTTACCAGATTTATGAGGCTTGAAGGATGTGTCGATG 3701

```

Db 3736 AGCTTATCTGAGTATTTAATTACTTACCAGATTATGAGCCCTGGAAGGTATGTCGATG 3795
QY 3702 GATAAATAGTATTTTAAAGAACTAATTCACATGAACCTAAATCATCAAGCAGCAGTG 3761
Db 3796 GATAAATAGTATTTTAAAGAACTAATTCACATGAACCTAAATCATCAAGCAGCAGTG 3855
QY 3762 GCCTCTACGTTTACTCCTTGTGCTGAAAAAAA 3794
Db 3856 GCCTCTACGTTTACTCCTTGTGAAAAAAA 3888

RESULT 13
AF342982
LOCUS Homo sapiens tankyrase 2 mRNA, complete cds.
DEFINITION AF342982
ACCESSION AF342982.1 GI:13430364
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3501)
Kaminker,P.G., Kim,S.H., Taylor,R.D., Zebajadian,Y., Funk,W.D.,
Morin,G.B., Yaswen,P. and Campisi,J.
TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
rapid induction of cell death upon overexpression
J. Biol. Chem. 276 (38), 35891-35899 (2001)
21443728
PUBMED 11454873
2 (bases 1 to 3501)
Kaminker,P.G., Campisi,J., Kim,S.H., Yaswen,P. and Morin,G.
Direct Submission
Submitted (25-JAN-2001) Life Sciences, Lawrence Berkeley National
Labs, 1 Cyclotron Rd 84-171, Berkeley, CA 94720, USA
FEATURES
source
1..3501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q23-q24"
1..3501
/note="TANK2; telomere-associated poly(ADP-ribose)
polymerase; contains ankyrin domain, sterile alpha motif
and PARP catalytic domain; similar to TNKS"
/codon_start=1
/product="tankyrase 2"
/protein_id="AK25811.1"
/db_xref="GI:13430365"
/translation="MSGRRCAGGGAACASAAAAEVEPAARELFEACRNGDVERVKRLV
TPEKNSRDTAGRKSTPLHFAAGFKRDVVEYLQNGANVOARDGGLIPLHNACSPG
HAEVNLLLRHGADPNARDNWNYPLEAAIKGKIDVCIVLHOCAPTIRNTDGPRA
LDADPSAKAVLTGEYKKDELESARSNEKMMALITPLNVNCHASDGRKSPHLHA
AGYNRPKIVOLLQHDGADVHAKDGLVPLHNACSYGHEVTELLVHRGACVNMDLW
LQAREADVTIRIKHLSLEMYNEKHPOTHEALHCAASPYPKRQICELLRKGAN
INTEKTEPLTHVASEKANDVVEVVKHEAKVNDNLNGOTSLHRAAYCGHLQTCR
LLSYGCDPNTIISLQGETALQMGNEVQQLQEGISLGNSEADRLLEAAKAGDVEV
KKLCTVQSVNCRDTEHQSTPLHFAAGYNRVVVEYLQHDADVHAKDGLVPLHNA
CSYGHVEAVELLVKGAVNVADLWKPTPLHEAAKAKYKIECKLLQLQHDGAPTKKND
GNTPLDLVKDGTDIQDLRGDAALAAKGLCLARKVKLSSPNVNCRTQGRHSPT
LHLAAGYNLEVAYLLOHGADVNAQDGGGLIPLHNAASYGHVDVAAKLYNACVNA
TDKWAFTPLHEAAGQRTQCALLAHAGDPTLKNQGGTPLDLVSADDDYSALLTAA
PPSALPCYKQVQLNGVRSPGATLADSSGSPSSLSAASSLNDLSGSPSELSSVVS
SSGTEGASSLEKKEKVPQVDESITQFVNLGLHLMIDIFEREQITLDVLVEMGHKELK
IGINAYGHRHKLKIGVERLISGQGLNPYTLNTSGSGTILIDLSPODKDEFQSVSEEM
QSTVREHRDGHGAGIENRYNLKIQVCKNKLWERYTHRRKEVSEENHNHMLF
HGSBFWNAIHKGDFERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHKDRS
CYCHROLLFCRVTLGKSFQFSAMKMAHSPPHSHSVTGRPSVNLALAEVVIYRGEQ
AYPELYTIQIMRPEGMVDG"
1044 a 707 c 840 g 909 t 1 others
BASE COUNT
ORIGIN

```

```

Query Match 86.9%; Score 3300.4; DB 9; Length 3501;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3301; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 403 AGSTTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAATGGTGAAGTGTCCAAGC 462
Db 198 AGSTTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAATGGTGAAGTGTCCAAGC 257
QY 463 ACCTGATGATGGGGCCCTTATTCCTTCATCAATCATGCTCTTTGGTTCATGCTGAAGT 522
Db 258 ACCTGATGATGGGGCCCTTATTCCTTCATCAATCATGCTCTTTGGTTCATGCTGAAGT 317
QY 523 AGTCAATCTCCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATAC 582
Db 318 AGTCAATCTCCTTTTGGGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATAC 377
QY 583 TCCTCTCATGAAGCTGCAATTTAAAGGAAAGATTCATCTTTGCAATTTGCTGTGTACAGCA 642
Db 378 TCCTCTCATGAAGCTGCAATTTAAAGGAAAGATTCATCTTTGCAATTTGCTGTGTACAGCA 437
QY 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 702
Db 438 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 497
QY 703 ATCTGCCAAAGCAGTCTTACTGCTGAATATAAGAAAGATGAATCTTTAGAAAAGTGCAG 762
Db 498 ATCTGCCAAAGCAGTCTTACTGCTGAATATAAGAAAGATGAATCTTTAGAAAAGTGCAG 557
QY 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAATGTCACACTGCCAGC 822
Db 558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAATGTCACACTGCCAGC 617
QY 823 AAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGATATAACAGAGTAAAGAT 882
Db 618 AAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGATATAACAGAGTAAAGAT 677
QY 883 TGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGTGATCTGGT 942
Db 678 TGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGTGATCTGGT 737
QY 943 ACCATTACACAATGCCTGCTTCTTATGGTCAATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 738 ACCATTACACAATGCCTGCTTCTTATGGTCAATTATGAAGTAACTGAACTTTTGGTCAAGCA 797
QY 1003 TGGTGCCTGTGTAATGCAATGGAGCTTGTGCAATTCACCTCTTCATGAGGAGCTTC 1062
Db 798 TGGTGCCTGTGTAATGCAATGGAGCTTGTGCAATTCACCTCTTCATGAGGAGCTTC 857
QY 1063 TAAGAACAGGCTTGAAGTATGCTCTCTCTTAAGTTATGGTGCAGACCCCAACACTGCT 1122
Db 858 TAAGAACAGGCTTGAAGTATGCTCTCTCTTAAGTTATGGTGCAGACCCCAACACTGCT 917
QY 1123 CAATTGTCAATAAAGTGCATATAGACTTTGGCTCCACACACAGTTAAAGAAAGATT 1182
Db 918 CAATTGTCAATAAAGTGCATATAGACTTTGGCTCCACACACAGTTAAAGAAAGATT 977
QY 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGAACTGATGTTACTCG 1242
Db 978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGAACTGATGTTACTCG 1037
QY 1243 AATCAAAAAACATCTCTCTGGAATTTTCAAGCATCCTCAACACATGAAC 1302
Db 1038 AATCAAAAAACATCTCTCTGGAATTTTCAAGCATCCTCAACACATGAAC 1097
QY 1303 AGCATTCATTTGCTGCTGCTATTCATCTCCATATCCCAAAAGAAAGCAATATGTGAACCTGT 1362
Db 1098 AGCATTCATTTGCTGCTGCTATTCATCTCCATATCCCAAAAGAAAGCAATATGTGAACCTGT 1157
QY 1363 GCTAAGAAAAAGGACCAACATCAATGAAAAGCTAAAGAAATTTCTGACTCCTCTGCACGT 1422
Db 1158 GCTAAGAAAAAGGACCAACATCAATGAAAAGCTAAAGAAATTTCTGACTCCTCTGCACGT 1217
QY 1423 GGCATCTGAGAAAGCTCATATATGATGTTGTTGAAGTAGTGTGTAACATGAACAAAGGT 1482

```



|||||  
Db 1218 GGCACTCAGAAAGCTCATATGATGTTGTTGAAGTACTGTGAAACATGAAGCAAGGT 1277  
|||||  
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTCTCTACAGAGAGCTGCATATTTGGTCACT 1542  
|||||  
Db 1278 TAATGCTCTGGATAATCTTGGTCAGACTCTCTACAGAGAGCTGCATATTTGGTCACT 1337  
|||||  
Qy 1543 ACAAACTGCCGCTACTCCCTGAGCTATGGGTGTCATCCTTAACATTTATATCCCTTCAGGG 1602  
|||||  
Db 1338 ACAAACTGCCGCTACTCCCTGAGCTATGGGTGTCATCCTTAACATTTATCCCTTCAGGG 1397  
|||||  
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTC 1662  
|||||  
Db 1398 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTC 1457  
|||||  
Qy 1663 ATTAGGTAAATCAGAGGAGACAGACAATGCTGTGAAGCTGCAAAAGCTGGAGATGTCGA 1722  
|||||  
Db 1458 ATTAGGTAAATCAGAGGACAGACAATGCTGTGAAGCTGCAAAAGCTGGAGATGTCGA 1517  
|||||  
Qy 1723 AACTGTAAAAAATCTGTACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCTCA 1782  
|||||  
Db 1518 AACTGTAAAAAATCTGTACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCTCA 1577  
|||||  
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCT 1842  
|||||  
Db 1578 GTCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCT 1637  
|||||  
Qy 1843 ACAGCATGGAGCTGATGTCATGCTTAAGATTAAGAGAGGCTGTACCTTTGCACAATGC 1902  
|||||  
Db 1638 ACAGCATGGAGCTGATGTCATGCTTAAGATTAAGAGAGGCTGTACCTTTGCACAATGC 1697  
|||||  
Qy 1903 ATGCTCTTATGGACATTTAAGAGTTCGCAACTCTCTGTAAACATGGAGCAGTAGTTAA 1962  
|||||  
Db 1698 ATGCTCTTATGGACATTTAAGAGTTCGCAACTCTCTGTAAACATGGAGCAGTAGTTAA 1757  
|||||  
Qy 1963 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA 2022  
|||||  
Db 1758 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA 1817  
|||||  
Qy 2023 AATTTCAAACTCTCTCCAGCATGGTCAGACCTTACCAAAAGGATGGAA 2082  
|||||  
Db 1818 AATTTGCAAACTCTCTCCAGCATGGTCAGACCTTACCAAAAGGATGGAA 1877  
|||||  
Qy 2083 TACTCCTTTGGATCTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142  
|||||  
Db 1878 TACTCCTTTGGATCTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 1937  
|||||  
Qy 2143 TGCAGCTTTGCTAGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTTGCTTTC 2202  
|||||  
Db 1938 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTTGCTTTC 1997  
|||||  
Qy 2203 TCCTGATAATGTAATTTGCCGCGATACCCAGGCAGACATTTCAACACCTTTACATTTAGC 2262  
|||||  
Db 1998 TCCTGATAATGTAATTTGCCGCGATACCCAGGCAGACATTTCAACACCTTTACATTTAGC 2057  
|||||  
Qy 2263 AGCTGTTTATAATTTAGAGTTGCAGAGTATTTGTTTACAACAGGAGCTGATGTGAA 2322  
|||||  
Db 2058 AGCTGTTTATAATTTAGAGTTGCAGAGTATTTGTTTACAACAGGAGCTGATGTGAA 2117  
|||||  
Qy 2323 TGCCCAAGACAAGGAGGACTTATCTTTTACATATGTCAGCATCTTACGGCATGTAGA 2382  
|||||  
Db 2118 TGCCCAAGACAAGGAGGACTTATCTTTTACATATGTCAGCATCTTACGGCATGTAGA 2177  
|||||  
Qy 2383 TGTAGCAGCTCTACTAATAAGTATATATGTCATGTCATGTCCTCAATGCCAGCACAATGGGCTTT 2442  
|||||  
Db 2178 TGTAGCAGCTCTACTAATAAGTATATATGTCATGTCCTCAATGCCAGCACAATGGGCTTT 2237  
|||||  
Qy 2443 CACACCTTTGCAGAGAGAGCCCAAAAGGAGCAGCAACAGCTTTGTGCTTTGTGCTAGC 2502  
|||||  
Db 2238 CACACCTTTGCAGAGAGAGCCCAAAAGGAGCAGCAACAGCTTTGTGCTTTGTGCTAGC 2297  
|||||  
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
|||||

Db 2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357  
|||||  
Qy 2563 AGCGATGATCTAGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622  
|||||  
Db 2358 AGCGATGATCTAGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417  
|||||  
Qy 2623 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTTC 2682  
|||||  
Db 2418 TTACAAGCCTCAAGTGTCTCAATGGTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTTC 2477  
|||||  
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTCTGACGCCAGCAGTCTTTCACAACTTATCTGG 2742  
|||||  
Db 2478 TTCAGGTCCATCTAGCCCATCAAGCCTTCTGACGCCAGCAGTCTTTCACAACTTATCTGG 2537  
|||||  
Qy 2743 GAGTTTTTTCAGAACTCTCTTTCAGTATGTTTCAAGTGAAGAGAGGGTCTTCCAGTTT 2802  
|||||  
Db 2538 GAGTTTTTTCAGAACTCTCTTTCAGTATGTTTCAAGTGAAGAGAGGGTCTTCCAGTTT 2597  
|||||  
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTACGATACTCAATTCGTGAAGAACTTTCG 2862  
|||||  
Db 2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTACGATACTCAATTCGTGAAGAACTTTCG 2657  
|||||  
Qy 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAGAAACAGATCACTTTTGGATGTATTAGTTGA 2922  
|||||  
Db 2658 ACTTGAGCACCTTAATGGATATATTTGAGAGAGAAACAGATCACTTTTGGATGTATTAGTTGA 2717  
|||||  
Qy 2923 GATGGGGCAACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGACATAGGCACAACCT 2982  
|||||  
Db 2718 GATGGGGCAACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGACATAGGCACAACCT 2777  
|||||  
Qy 2983 AATTAAAGAGTCGAGAGACTTATCTCCGAGACAACAGGCTTAAACCATATATTAACTTT 3042  
|||||  
Db 2778 AATTAAAGAGTCGAGAGACTTATCTCCGAGACAACAGGCTTAAACCATATATTAACTTT 2837  
|||||  
Qy 3043 GAACACTCTGCTGAGTGAACAACTTCTTATAGATCTGCTCTCTGATGATAAAGAGTTTCA 3102  
|||||  
Db 2838 GAACACTCTGCTGAGTGAACAACTTCTTATAGATCTGCTCTCTGATGATAAAGAGTTTCA 2897  
|||||  
Qy 3103 GTCGTGGAGGAGAGATGCAAAAGTACAGTTCGAGAGCAGAGATGGAGGTCTATCGAG 3162  
|||||  
Db 2898 GTCGTGGAGGAGAGATGCAAAAGTACAGTTCGAGAGCAGAGATGGAGGTCTATCGAG 2957  
|||||  
Qy 3163 TGAATCTTCAACAGATACAAATATCTCAAGATTCAGAAGTTTGAACAAAGAACTATG 3222  
|||||  
Db 2958 TGAATCTTCAACAGATACAAATATCTCAAGATTCAGAAGTTTGAACAAAGAACTATG 3017  
|||||  
Qy 3223 GGAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATGA 3282  
|||||  
Db 3018 GGAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCAATGCCAATGA 3077  
|||||  
Qy 3283 ACGAATGCTATTTTCATGGGCTCTCTTTTGTGAATGCAATTTATCCACAAAGGCTTTGATGA 3342  
|||||  
Db 3078 ACGAATGCTATTTTCATGGGCTCTCTTTTGTGAATGCAATTTATCCACAAAGGCTTTGATGA 3137  
|||||  
Qy 3343 AAGCATGCTTACATAGTGGTGTGTTTGGAGCTGGCATTTATTTTGTCTGAAAACCTCTTC 3402  
|||||  
Db 3138 AAGCATGCTTACATAGTGGTGTGTTTGGAGCTGGCATTTATTTTGTCTGAAAACCTCTTC 3197  
|||||  
Qy 3403 CAAAAGCAATCAATATGATATATGGAATTTGGAGGAGTACTGGGGTGTCCAGTTTCAAAAGA 3462  
|||||  
Db 3198 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTGTCCAGTTTCAAAAGA 3257  
|||||  
Qy 3463 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGGGTAACTTTGGGAAAGTC 3522  
|||||  
Db 3258 CAGATCTTGTATCTTTGCCACAGGAGCTGCTCTTTTGGGGTAACTTTGGGAAAGTC 3317  
|||||  
Qy 3523 TTTCTCTCAGTTTCACTGCAATGAAATGGCACAATTCCTCCAGGTCATCACTCAGTCAAC 3582  
|||||  
Db 3318 TTTCTCTCAGTTTCACTGCAATGAAATGGCACAATTCCTCCAGGTCATCACTCAGTCAAC 3377  
|||||  
Qy 3583 TGGTAGGCCAGTGTAATGGCCTTAGCATTAGCTGGAATATGTTATTTACAGAGAGAAACA 3642  
|||||  
Db 3378 TGGTAGGCCAGTGTAATGGCCTTAGCATTAGCTGGAATATGTTATTTACAGAGAGAAACA 3437  
|||||



QY 3643 GGCTTACCTCAGTATTTAATTACTTACCAGATTATGAGGCTGAAGGTATGTCGATGG 3702  
|||||  
Db 3438 GGCTTACCTCAGTATTTAATTACTTACCAGATTATGAGGCTGAAGGTATGTCGATGG 3497  
|||||

QY 3703 ATAA 3706  
|||  
Db 3498 ATAA 3501  
|||

RESULT 14  
AX029397 LOCUS 3400 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 1 from Patent WO9915647.  
ACCESSION AX029397  
VERSION AX029397.1 GI:10190194  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3400)  
AUTHORS Sutherland, R.L. and Daly, R.J.  
TITLE A potential effector for the grb7 family of signalling proteins  
JOURNAL Patent: WO 9915647-A 1 01-APR-1999;  
GARVAN INST MED RES (AU) ; SUTHERLAND ROBERT LYNDASAY (AU) ; DALY  
ROGER JOHN (AU)  
FEATURES  
source Location/Qualifiers  
1..3400  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1059 a 665 c 760 g 916 t  
ORIGIN

Query Match 86.8%; Score 3297; DB 6; Length 3400;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 482 ATTCTCTTCATATGATGCTCTTTGGTTCATGCTGAAGTAGTCAATCTCCTTTTCGGA 541  
|||||  
Db 1 ATTCTCTTCATATGATGCTCTTTGGTTCATGCTGAAGTAGTCAATCTCCTTTTCGGA 60  
|||||

QY 542 CATGGTCGAGACCCCAATGCTCGAGATAATTTGGAATTATATCTCTCTCCATGAAGCTGCA 601  
|||||  
Db 61 CATGGTCGAGACCCCAATGCTCGAGATAATTTGGAATTATATCTCTCTCCATGAAGCTGCA 120  
|||||

QY 602 ATTAAGGAAGAATGATGTTGGATGCTGTTACAGATGAGCTGAGCCCAACCATC 661  
|||||  
Db 121 ATTAAGGAAGAATGATGTTGGATGCTGTTACAGATGAGCTGAGCCCAACCATC 180  
|||||

QY 662 CGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAGCAGTGCTT 721  
|||||  
Db 181 CGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAGCAGTGCTT 240  
|||||

QY 722 ACTGCTGAATATAAGAAAGATGAATCTCTTAGAAAGTGCCAGAGTGGCAATGAAGAAAAA 781  
|||||  
Db 241 ACTGCTGAATATAAGAAAGATGAATCTCTTAGAAAGTGCCAGAGTGGCAATGAAGAAAAA 300  
|||||

QY 782 ATGATGGCTCTACTCACACCATTAATGTCACCTGCCACGCAAGTGATGGCAGAAAGTCA 841  
|||||  
Db 301 ATGATGGCTCTACTCACACCATTAATGTCACCTGCCACGCAAGTGATGGCAGAAAGTCA 360  
|||||

QY 842 ACTCCATTACATTTGGCAGCAGGATATAACAGATGAAGATTGTACAGCTGTTACTGCAA 901  
|||||  
Db 361 ACTCCATTACATTTGGCAGCAGGATATAACAGATGAAGATTGTACAGCTGTTACTGCAA 420  
|||||

QY 902 CATGGAGCTGATGTCATGCTAAAGATAAAGTGATCTGTGTTACCATTACCAATGCCCTGT 961  
|||||  
Db 421 CATGGAGCTGATGTCATGCTAAAGATAAAGTGATCTGTGTTACCATTACCAATGCCCTGT 480  
|||||

QY 962 TCTTATGGTCATATTGAAGTACTGAACCTTTTGGTCAAGCATGGTGCCTGTGTAATGCA 1021  
|||||  
Db 481 TCTTATGGTCATATTGAAGTACTGAACCTTTTGGTCAAGCATGGTGCCTGTGTAATGCA 1620  
|||||

QY 1022 ATGGACTTGTGGCAATTCACCTCCTCTTCATGAGGAGCTTCTTAAGAACAGGCTTGAAGTA 1081  
|||||  
Db 541 ATGGACTTGTGGCAATTCACCTCCTCTTCATGAGGAGCTTCTTAAGAACAGGCTTGAAGTA 600  
|||||

QY 1082 TGTTCCTCTCTCTTAAGTTATGTCGAGACCCCAACACTGCTCAATTTGTACCAATAAAAAT 1141  
|||||  
Db 601 TGTTCCTCTCTCTTAAGTTATGTCGAGACCCCAACACTGCTCAATTTGTAAAGTAAAAAT 660  
|||||

QY 1142 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 1201  
|||||  
Db 661 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 720  
|||||

QY 1202 CACTCTGCTGTCGAAGCTGCACGAGAAGCTGATGTTTACTCGAATCAAAAACATCTCTCT 1261  
|||||  
Db 721 CACTCTGCTGTCGAAGCTGCACGAGAAGCTGATGTTTACTCGAATCAAAAACATCTCTCT 780  
|||||

QY 1262 CTGGAATGCTGAATTTCAAGCATCTCTCAACACATGAACACAGCATTTGCATTTGTGCTGT 1321  
|||||  
Db 781 CTGGAATGCTGAATTTCAAGCATCTCTCAACACATGAACACAGCATTTGCATTTGTGCTGT 840  
|||||

QY 1322 GCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAATCTGTTGCTRAGAAAAGGACCAAC 1381  
|||||  
Db 841 GCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAATCTGTTGCTAAGAAAAGGACCAAC 900  
|||||

QY 1382 ATCAATGAAAAGACTAAAGAAATCTTGTACTCTCTGCAGTGGCATCTGAGAAAAGCTCAT 1441  
|||||  
Db 901 ATCAATGAAAAGACTAAAGAAATCTTGTACTCTCTGCAGTGGCATCTGAGAAAAGCTCAT 960  
|||||

QY 1442 AATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGGTTAATGCTCTGGATATATCTT 1501  
|||||  
Db 961 AATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGGTTAATGCTCTGGATATATCTT 1020  
|||||

QY 1502 GGTACACATCTCTACACAGCTGCATATTTGTTGGTCACTCTACAAACCTTGGCCCTACTC 1561  
|||||  
Db 1021 GGTACACATCTCTACACAGCTGCATATTTGTTGGTCACTCTACAAACCTTGGCCCTACTC 1080  
|||||

QY 1562 CTGAGCTATGGGTGTGATCCTAACATTAATCCTTTCAGGGCTTTTACTGCTTTTACAGATG 1621  
|||||  
Db 1081 CTGAGCTATGGGTGTGATCCTAACATTAATCCTTTCAGGGCTTTTACTGCTTTTACAGATG 1140  
|||||

QY 1622 GGAATGAAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATAGGTATTTTACAGAGGCA 1681  
|||||  
Db 1141 GGAATGAAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATTTAGGTATTTTACAGAGGCA 1200  
|||||

QY 1682 GACAGCAATTTGCTGCAAGCTGCAAGGCTGAGATGTCGAAACTGTAAAAAACTGTGT 1741  
|||||  
Db 1201 GACAGCAATTTGCTGCAAGCTGCAAGGCTGAGATGTCGAAACTGTAAAAAACTGTGT 1260  
|||||

QY 1742 ACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCCTTCATTTT 1801  
|||||  
Db 1261 ACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCCTTCATTTT 1320  
|||||

QY 1802 GCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCTACAGATGGAGCTGATGTG 1861  
|||||  
Db 1321 GCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCTACAGATGGAGCTGATGTG 1380  
|||||

QY 1862 CATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTTAT 1921  
|||||  
Db 1381 CATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTTAT 1440  
|||||

QY 1922 GAAGTTTCAGAACTTCTTTTAAACATGAGCAGTAGTAAATAGTGTATTTATGGAAA 1981  
|||||  
Db 1441 GAAGTTTCAGAACTTCTTTTAAACATGAGCAGTAGTAAATAGTGTATTTATGGAAA 1500  
|||||

QY 1982 TTTACACCTTTTACATGAAGCAGCAGCAAGAAAATATGAAAATTTGCAAACTTCTGCTC 2041  
|||||  
Db 1501 TTTACACCTTTTACATGAAGCAGCAGCAAGAAAATATGAAAATTTGCAAACTTCTGCTC 1560  
|||||

QY 2042 CAGCATGGTCAGACCCCTACCAAAAACACAGGATGGAATACTCTTTGGATCTTGT 2101  
|||||  
Db 1561 CAGCATGGTCAGACCCCTACCAAAAACACAGGATGGAATACTCTTTGGATCTTGT 1620  
|||||



HLAAGYNRVIVQLLQHGADVHAKOKGDLVPLHNACSYGHVEYVTELLVKGACVAM  
 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSADILAPTPOKLERLAEYFK  
 GHSILOAREADVTIRIKKHLSEWAFKHPQTHETALHCAASPYKPKQICELLRRK  
 GANINEXTKELPLPHVASEKANDVVEVYVKEAKVNALDNLGQTSLHRAATCGHLQ  
 TCRLLSYGDPNIIISLQGTALOMGNENVOQLLQEGISLGNSEADQRDLLEAAKAGDV  
 ETVKLLCTGVSNCRLEGRQSTPLHFAAGYNRVYVEYLLOHGADVHAKDKGLVPL  
 HNACSYGHEVAELLVKGAVVNVADLWKFPLHEAAKGYEICKLQLLQHGADPTKK  
 NRQNTPLDLVKGDDTDIQLLRGDAALLDAARKGLARVKLSSPDNVNCRDTCGRH  
 STPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIYNAC  
 VNATDKWAFPLHEAAKGRTOQLCALLLHAGADPTLKNQEGPTPLDLVSDVVSALLT  
 AAMPSPALPCYPOVLNGRSCATADALUSSPSSLSAASSLNDLSGSELUSS  
 VYSSSGTGASSLEKKEKPGVDFSIQFVRNIGLEHMDIFEREQITLDLVPMGHE  
 LKEITINAYGHRKILKGVRLISGQQLNPLTLATSGSTLILDLSPDKRFQSV  
 EEMOSTVREHRDGHAGGIFNRYNLIKQKCNKWLERYTHRRKEVSEENHNHNER  
 MLPHGSPFVNAI IHKGDFERHAYIGGMFGAGIYFAENS SKSNQYVVGIGGTCPCVHK  
 DRSCYICHRQLLFCRVPLKSFQFSAMKMAHSPPHSHSVTGRPSVNLALAEYIYR  
 GEQAYPELTIYQIMRPEGMVDG\*

BASE COUNT 1046 a 710 c 843 g 909 t

Query Match	86.8%	Score 3296.2	DB 6	Length 3508
Best Local Similarity	99.9%	Pred. No. 0		
Matches 3298	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY 403	AGGTTTGGCGGAAGACGTAGTTGAATTTGCTTCAGAAATGGTCAAGTGTCCAAAGC	462		
DB 208	AGGTTTGGCGGAAGACGTAGTTGAATTTGCTTCAGAAATGGTCAAGTGTCCAAAGC	267		
QY 463	ACGTGATGATGGGGCCCTTATTCCTTCATTAATGATGCTCTTTTGGTCATGCTGAAGT	522		
DB 268	ACGTGATGATGGGGCCCTTATTCCTTCATTAATGATGCTCTTTTGGTCATGCTGAAGT	327		
QY 523	AGTCAATCTCTTTTGGCAGATGTGCAGACCCCAATCCTCGAGATAATTCGAATTTATAC	582		
DB 328	AGTCAATCTCTTTTGGCAGATGTGCAGACCCCAATCCTCGAGATAATTCGAATTTATAC	387		
QY 583	TCCTCTCCATGAAGTGCATTAATAAGGAAAGATTGATGTTTGCATTTGCTGTTTACAGCA	642		
DB 388	TCCTCTCCATGAAGTGCATTAATAAGGAAAGATTGATGTTTGCATTTGCTGTTTACAGCA	447		
QY 643	TGGAGCTGAGCCAAACATCCGAAATACAGATGGAGGACAGCATTTGGATTTAGCAGATCC	702		
DB 448	TGGAGCTGAGCCAAACATCCGAAATACAGATGGAGGACAGCATTTGGATTTAGCAGATCC	507		
QY 703	ATCTGCAAGACAGTCTACTGTTGTAATATAAGAAAGATGAATCTTAGAAAGTGCCAG	762		
DB 508	ATCTGCAAGACAGTCTACTGTTGTAATATAAGAAAGATGAATCTTAGAAAGTGCCAG	567		
QY 763	GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTAATGATCAACTGCCACGC	822		
DB 568	GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTAATGATCAACTGCCACGC	627		
QY 823	AGTGTATGCGCAAGATCACTCCATTTACATTTGGCAGCAGATATAACAGATGAAGAT	882		
DB 628	AGTGTATGCGCAAGATCACTCCATTTACATTTGGCAGCAGATATAACAGATGAAGAT	687		
QY 883	TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGGT	942		
DB 688	TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGGT	747		
QY 943	ACCATTACCAATAGCTGCTTCTTATGGTCATTTAAGTAACTGAACCTTTTGGTCAAGCA	1002		
DB 748	ACCATTACCAATAGCTGCTTCTTATGGTCATTTAAGTAACTGAACCTTTTGGTCAAGCA	807		
QY 1003	TGGTGCCTGTGAATGCAATGGACTTTGGCAATTTCACTCTTTCATGAGGACGCTTC	1062		
DB 808	TGGTGCCTGTGAATGCAATGGACTTTGGCAATTTCACTCTTTCATGAGGACGCTTC	867		
QY 1063	TAAAGACAGGGTTGAAGTATGTTCTCTCTTCTTAAGTATGGTCAGACCCCAACACTGCT	1122		
DB 868	TAAAGACAGGGTTGAAGTATGTTCTCTCTTCTTAAGTATGGTCAGACCCCAACACTGCT	927		
QY 1123	CAATTGTCAATAAAGTGTCTATAGACTTTGGCTCCCAACACCACTTTAAAGAAAGATT	1182		

DB 928	CAATTGTGCACAAATAAAGTGCTATAGACTTGGCTCCACACCACTGTTAAAGAAAAGATT	987
QY 1183	AGCATATGAATTTAAAGGCCACTCGTGTGCAAGCTGCACGAGAGCTGATGTTACTCG	1242
DB 988	AGCATATGAATTTAAAGGCCACTCGTGTGCAAGCTGCACGAGAGCTGATGTTACTCG	1047
QY 1243	AATCAAAAAACATCTCTCTCTGGAATTTCAAGCATCTCTCAACACATGAAC	1302
DB 1048	AATCAAAAAACATCTCTCTGGAATTTCAAGCATCTCTCAACACATGAAC	1107
QY 1303	AGCATATGATTTGCTGCTGCATCTCCATATCCCAAAAGAAACAATAATGTGAATGTT	1362
DB 1108	AGCATATGATTTGCTGCTGCATCTCCATATCCCAAAAGAAACAATAATGTGAATGTT	1167
QY 1363	GCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAAGAAATTTTGAACCTCTGACGCT	1422
DB 1168	GCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAAGAAATTTTGAACCTCTGACGCT	1227
QY 1423	GGCATCTGAGAAAAGCTCATATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT	1482
DB 1228	GGCATCTGAGAAAAGCTCATATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT	1287
QY 1483	TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTTGCTCATCT	1542
DB 1288	TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGTTGCTCATCT	1347
QY 1543	ACAAACCTGCGGCTACTCTGAGCTATGGTGTGATCTTAAACATATATATCCCTTCAGG	1602
DB 1348	ACAAACCTGCGGCTACTCTGAGCTATGGTGTGATCTTAAACATATATATCCCTTCAGG	1407
QY 1603	CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGCTATCTC	1662
DB 1408	CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGCTATCTC	1467
QY 1663	ATTAGGTAATTCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAAAGCTGGAGATGCTGA	1722
DB 1468	ATTAGGTAATTCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAAAGCTGGAGATGCTGA	1527
QY 1723	AACGTGTAATAAAGCTGTTACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTCA	1782
DB 1528	AACGTGTAATAAAGCTGTTACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTCA	1587
QY 1783	GTCTACACCACTTCATTTTCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT	1842
DB 1588	GTCTACACCACTTCATTTTCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT	1647
QY 1843	ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACCTTTTGCACATGC	1902
DB 1648	ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACCTTTTGCACATGC	1707
QY 1903	ATGTTCTTATGGACATTTAAGTGTGAGACTTCTTGTAAACATGAGCAGCTAGTGTAA	1962
DB 1708	ATGTTCTTATGGACATTTAAGTGTGAGACTTCTTGTAAACATGAGCAGCTAGTGTAA	1767
QY 1963	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	2022
DB 1768	TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA	1827
QY 2023	AATTTGCAAACTTCTGCTCCAGCATGTCAGAGCCCTTACCAAAAAAACAAGGATGGA	2082
DB 1828	AATTTGCAAACTTCTGCTCCAGCATGTCAGAGCCCTTACCAAAAAAACAAGGATGGA	1887
QY 2083	TACTCTCTTGGATCTCTTAAAGATGAGATACAGATATTTCAAGATCTGCTTGGGAGGA	2142
DB 1888	TACTCTCTTGGATCTCTTAAAGATGAGATACAGATATTTCAAGATCTGCTTGGGAGGA	1947
QY 2143	TGCAGCTTTGCTAGATGCTCCAAAGAGGGTTGTTTACCCAGAGTGAAGATTTGCTTTC	2202
DB 1948	TGCAGCTTTGCTAGATGCTCCAAAGAGGGTTGTTTACCCAGAGTGAAGATTTGCTTTC	2007
QY 2203	TCCTGATAATGTAATTTGCCGCGATACCCAGGAGACATTTCAACACCTTTTACATTTAGC	2262

Db 2008 TCCTGATAATGTAAATGCCCGGATACCCAAAGGCGAGACATCAACACCCCTTTACATTTAGC 2067  
QY 2263 AGCTGGTTTATAATATTAGAAAGTTGCAGAGTATTGTTTACAACACGAGCTGTAGTGAA 2322  
Db 2068 AGCTGGTTTATAATATTAGAAAGTTGCAGAGTATTGTTTACAACACGAGCTGTAGTGAA 2127  
QY 2323 TGCCCAAGACAAAGAGGAGCTTATPCCCTTTACATTAATGCAGCATCTTACGGGCATGTAGA 2382  
Db 2128 TGCCCAAGACAAAGAGGAGCTTATPCCCTTTACATTAATGCAGCATCTTACGGGCATGTAGA 2187  
QY 2393 TGTACAGCTCTACTAATAAAGTATAATGCATGTCTCAATGCCACGGACAAATGGCTTTT 2442  
Db 2188 TGTACAGCTCTACTAATAAAGTATAATGCATGTCTCAATGCCACGGACAAATGGCTTTT 2247  
QY 2443 CACACCTTTTGACGAAGACAGCCCAAAAGGACGAACACAGCTTTTGTCTTTGCTTACG 2502  
Db 2248 CACACCTTTTGACGAAGACAGCCCAAAAGGACGAACACAGCTTTTGTCTTTGCTTACG 2307  
QY 2503 CCATGGAGCTGACCCGACTCTTTAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
Db 2308 CCATGGAGCTGACCCGACTCTTTAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367  
QY 2563 AGCGGATGATGTCAGGCTCTCTCACAGCAGCCATGCCCATCTGCTCTGCCCTCTTG 2622  
Db 2368 AGCAGATGATGTCAGGCTCTCTCACAGCAGCCATGCCCATCTGCTCTGCCCTCTTG 2427  
QY 2623 TTACAAGCTCAAGTCTCAATGGTGTGAGAGCCGAGGACCTGCGAGATGCTCTCTC 2682  
Db 2428 TTACAAGCTCAAGTCTCAATGGTGTGAGAGCCGAGGACCTGCGAGATGCTCTCTC 2487  
QY 2683 TTCAGGTCCATCTAGCCCATCAAGCTTTCTCGACGCCAGCAGCTTTGACAACTTATCTGG 2742  
Db 2488 TTCAGGTCCATCTAGCCCATCAAGCTTTCTCGACGCCAGCAGCTTTGACAACTTATCTGG 2547  
QY 2743 GAGTWTTCAGAACCTCTCTCAGTAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2802  
Db 2548 GAGTWTTCAGAACCTCTCTCAGTAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 2607  
QY 2803 GGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862  
Db 2608 GCAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667  
QY 2863 ACTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922  
Db 2668 ACTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727  
QY 2923 GATGGGGCAAGGAGCTGAAGGAGATGGAAATCAATGCTTATGGACATAGGCACAAACT 2982  
Db 2728 GATGGGGCAAGGAGCTGAAGGAGATGGAAATCAATGCTTATGGACATAGGCACAAACT 2787  
QY 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAAAGGCTTTAACCCATATTTAACTTT 3042  
Db 2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAAAGGCTTTAACCCATATTTAACTTT 2847  
QY 3043 GAACACCTCTGTTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATAAGAGATTCA 3102  
Db 2848 GAACACCTCTGTTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATAAGAGATTCA 2907  
QY 3103 GTCTGTGGAGGAGATGCAAGTACAGTTCGAGAGCAGAGATGGAGGTGATGAGG 3162  
Db 2908 GTCTGTGGAGGAGATGCAAGTACAGTTCGAGAGCAGAGATGGAGGTGATGAGG 2967  
QY 3163 TGGAACTCTTCAACAGATACAATATTCTCAAGATTTCAGAAAGTTTGTAAACAAAGAACTATG 3222  
Db 2968 TGGAACTCTTCAACAGATACAATATTCTCAAGATTTCAGAAAGTTTGTAAACAAAGAACTATG 3027  
QY 3223 GGAAGATACACTACCCGAGAGAAAGAGTTTCTGAAGAAAACCAACCATGCCAATGA 3282  
Db 3028 GGAAGATACACTACCCGAGAGAAAGAGTTTCTGAAGAAAACCAACCATGCCAATGA 3087  
QY 3283 ACGAATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTTATCCACAAAGGCTTTGATGA 3342  
Db 3088 ACGAATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTTATCCACAAAGGCTTTGATGA 3147

QY 3343 AAGGCATCGGTACATAGGTGGTATATGTTTGGAGCTGGCATTTATTTTGTGAAAACTCTTC 3402  
Db 3148 AAGGCATCGGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGTGAAAACTCTTC 3207  
QY 3403 CAAAAGCAATCAATATGTATATGGAATTTGGAGGAGTACTGGGTCTCCAGTTTCACAAAGA 3462  
Db 3208 CAAAAGCAATCAATATGTATATGGAATTTGGAGGAGTACTGGGTCTCCAGTTTCACAAAGA 3267  
QY 3463 CAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTCCCGGTAACCTTGGGAAAGTC 3522  
Db 3268 CAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTCCCGGTAACCTTGGGAAAGTC 3327  
QY 3523 TTTCTCTGAGTTTCAAGTCAATGAAAAATGGCAATTTCTCTCCAGGTCACTCAGTCACTC 3582  
Db 3328 TTTCTCTGAGTTTCAAGTCAATGAAAAATGGCAATTTCTCTCCAGGTCACTCAGTCACTC 3387  
QY 3583 TGGTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACA 3642  
Db 3388 TGGTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACA 3447  
QY 3643 GGCTTATCTGAGTATTTAAATTTACTTTACCAGATTATGAGGCCTGAAGGTATGTCGATGG 3702  
Db 3448 GGCTTATCTGAGTATTTAAATTTACTTTACCAGATTATGAGGCCTGAAGGTATGTCGATGG 3507  
QY 3703 A 3703  
Db 3508 A 3508

Search completed: February 11, 2003, 18:25:58  
Job time : 9476.36 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:17:23 ; Search time 717.206 Seconds  
(without alignments)  
11922.436 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 cttgaagacactggtttc.....ccttgctgaaaaa 3797

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID2/gcgdata/geneq/geneq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneq/geneq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneq/geneq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneq/geneq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneq/geneq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneq/geneq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneq/geneq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneq/geneq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneq/geneq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneq/geneq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneq/geneq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneq/geneq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneq/geneq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneq/geneq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneq/geneq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneq/geneq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneq/geneq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneq/geneq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneq/geneq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneq/geneq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneq/geneq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneq/geneq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneq/geneq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneq/geneq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3797	100.0	3797	22	AAA91487
2	3393.4	89.4	3816	22	AAA91488
3	3388.8	89.2	4275	21	AAC56825
4	3388.8	89.2	6019	22	AAC85294
5	3387.2	89.2	4127	22	AAF63926
6	3387.2	89.2	4406	22	AAF63930
7	3387.2	89.2	4992	22	AAF63953
8	3387.2	89.2	5002	22	AAF63952
9	3382.8	89.1	4512	22	AAD02578

10	3297	86.8	3400	20	AA25366	Human Grb7 effecto
11	3296.2	86.8	3508	22	AAF63837	Human tankyrase2 r
12	2969.6	78.2	4493	21	AAC66823	Human tankyrase II
13	2968.2	78.2	4296	21	AAC66824	Human tankyrase II
14	2774.6	73.1	5005	22	AAC89706	Human adult T-cell
15	2439.8	64.3	2971	22	AAF63919	Human tankyrase2 c
16	2439.8	64.3	3253	22	AAF63917	Human tankyrase2 c
17	2357.8	62.1	3799	22	AAF63918	Human tankyrase2 c
18	1813.4	47.8	3984	22	AAF63838	Human tankyrase I
19	1813.4	47.8	4134	21	AAC66826	Human tankyrase I
20	1813.4	47.8	4134	21	AAC66827	Human tankyrase I
21	1813.4	47.8	4134	24	ABK48629	Human tankyrase CD
22	1794.8	47.3	3254	22	AAC85295	Human tankyrase CD
23	1449.6	38.2	4491	21	AAZ29632	Mouse SPANK CDNA.
24	1449.6	38.2	4491	24	ABK48634	Human truncated ta
25	1359.4	35.8	1568	22	AAF63925	Human tankyrase2 c
26	1297.8	34.2	4657	21	AAZ29633	Human truncated ta
27	1297.8	34.2	4657	24	ABK48635	Human truncated ta
28	1240.4	32.7	1466	22	ABA08564	Human Grb7 effecto
29	1172.2	30.9	1242	22	AAC89701	Human adult T-cell
30	1168.4	30.8	3308	22	AAF63990	Human tankyrase2 e
31	1084.8	28.6	1103	22	AAF63973	Human TRF1 TANK2 b
32	1076.8	28.4	3006	22	AAH16624	Human cDNA sequenc
33	1008.8	26.6	5482	22	AAF63955	Drosophila tankyra
34	1008.8	26.6	5484	23	ABL08155	Drosophila melanog
35	994	26.2	2409	22	AAK99406	DNA of APP related
36	990	26.1	1007	22	AAC89703	Human adult T-cell
37	786	20.7	1344	21	AAC77578	Human ORFX ORF3133
38	784.2	20.7	838	22	AAH08231	Human cDNA clone (
39	748.6	19.7	10194	23	ABL08154	Drosophila melanog
40	574	15.1	582	22	AAF63875	Human tankyrase1 r
41	560	14.7	2793	21	AAC79928	Human secreted pro
42	518.4	13.7	520	21	AAC66828	Ankyrin-related AD
43	459	12.1	460	22	AAH35131	Human colon cancer
44	459	12.1	460	22	AAF63839	Human tankyrase1 r
45	396.8	10.5	402	22	AAH42791	Nucleotide sequenc

#### ALIGNMENTS

#### RESULT 1

AAA91487  
ID AAA91487 standard; DNA; 3797 BP.

XX AAA91487;

XX 06-AUG-2001 (first entry)

XX Tankyrase homologue isotype 1 coding sequence.

XX Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;  
XX cell cycle protein; cell cycle associated disorder; cancer; gene mapping;  
XX chromosome mapping; gene therapy; vaccine; ds.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 404..3706

XX /\*tag= a

XX /partial

XX /product= "TaHo1"

XX /note= "Tankyrase homologue isotype 1; No start

XX codon given"

XX WO200130987-A2.

XX PD 03-MAY-2001.

XX XX 25-OCT-2000; 2000WO-US41528.

XX XX 25-OCT-1999; 99US-0427154.

(RIGE-) RIGEL PHARM INC.  
Luo Y, Chan E, Xu X, Huang B;  
WPI: 2001-300503/31.  
P-PSDB: AAY97748.  
Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing or preventing cell proliferation in cells, and for diagnosing, treating or preventing cell cycle associated disorders such as cancer.  
Claim 13; Fig 1; 63pp; English.  
This sequence encodes the Tankyrase homologue isotype 1 (TaHo-1) protein of the invention. The invention also relates to the TaHo-2 protein. The TaHo proteins are useful for inducing or preventing cell proliferation in cells, and in the study or treatment of conditions mediated by the cell cycle proteins, such as to diagnose, treat or prevent cell cycle associated disorders, preferably cancer. The TaHo coding sequences are useful as hybridisation probes, in chromosome and gene mapping and in the generation of anti-sense DNA and RNA. The coding sequences are also useful for the preparation of TaHo, for generating either transgenic animals or knock out animals which, in turn, are useful in a development and screening of therapeutically useful agents, in gene therapy, as vaccine, and for construction of hybridisation probes for mapping the gene which encodes TaHo and for the genetic analysis of individuals with genetic disorders. The TaHo proteins, and their coding sequences are useful in screening assays.  
Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 other;  
Query Match 100.0%; Score 3797; DB 22; Length 3797;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGAAGACACTGGATTTCATACCTTTTGGCTGGGGTTATCTCTCTGTGCTCACTACAT 60  
DB 1 CTTTGAAGACACTGGATTTCATACCTTTTGGCTGGGGTTATCTCTCTGTGCTCACTACAT 60  
QY 61 AGACAAATATTAGCTGTGAGCAGATCTTTTGGTGGTCTTCTGTAGTCCCGAGTTAG 120  
DB 61 AGACAAATATTAGCTGTGAGCAGATCTTTTGGTGGTCTTCTGTAGTCCCGAGTTAG 120  
QY 121 CAGAAACATCTGAGATAGATCTGGGAAAGAAATCTAGCAAGAGTTTGTGCACGTGA 180  
DB 121 CAGAAACATCTGAGATAGATCTGGGAAAGAAATCTAGCAAGAGTTTGTGCACGTGA 180  
QY 181 TCATAAGSTTGTGATTTACATATTAAAGTTTATCTTTGAACATCTGAAATGTATACA 240  
DB 181 TCATAAGSTTGTGATTTACATATTAAAGTTTATCTTTGAACATCTGAAATGTATACA 240  
QY 241 TACTAAATATCAGACACTCTATTGTAGAGTCAGAAACATTTGAACCTTTGAGCTTTCAGTC 300  
DB 241 TACTAAATATCAGAACTCTATTGTAGAGTCAGAAACATTTGAACCTTTGAGCTTTCAGTC 300  
QY 301 ACTTATTTGTATCTTTCTTTAGGTTAGCAGTAGTACCACCCAGCAGCTGCTTAGGT 360  
DB 301 ACTTATTTGTATCTTTCTTTAGGTTAGCAGTAGTACCACCCAGCAGCTGCTTAGGT 360  
QY 361 ACCACTGCTGTAGTGAGAGTCCTCTGGCTTTATCAATTAAGSTTTTGGCGGAAAGA 420  
DB 361 ACCACTGCTGTAGTGAGAGTCCTCTGGCTTTATCAATTAAGSTTTTGGCGGAAAGA 420  
QY 421 CGTAGTTGAATATTTGCTTTCAGAAATGGTGAAGTGTGCAAGCAGCTGATGATGGGGCT 480  
DB 421 CGTAGTTGAATATTTGCTTTCAGAAATGGTGAAGTGTGCAAGCAGCTGATGATGGGGCT 480  
QY 481 TATTCTCTTCAATATGATGCTCTTTTGGTCAATGCTGAGTAGTCAATCTCTTTGGG 540  
DB 481 TATTCTCTTCAATATGATGCTCTTTTGGTCAATGCTGAGTAGTCAATCTCTTTGGG 540  
QY 541 ACATGGTCAGACCCCAATGCTCGAGATAATTTGGATTTATCTCTCTCCATCAAGCTGC 600

DB 541 ACATGGTCAGACCCCAATGCTCGAGATAATTTGGAATTTACTCTCTCCATCAAGCTGC 600  
QY 601 AATTAAAGAAAGATTTGATTTTGGCTTTGCTGTTTACAGCATGGAGCTGAGCAACCAT 660  
DB 601 AATTAAAGAAAGATTTGATTTTGGCTTTGCTGTTTACAGCATGGAGCTGAGCAACCAT 660  
QY 661 CCGAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGT 720  
DB 661 CCGAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGT 720  
QY 721 TACTGCTGAATATPAGAAAGATGAACTCTTAGAAAGTGGCAGAGTGGCAATCAAGAAA 780  
DB 721 TACTGCTGAATATPAGAAAGATGAACTCTTAGAAAGTGGCAGAGTGGCAATCAAGAAA 780  
QY 781 AATGATGGCTTACTCACACCATTTAAATGTCACTGCCAGCAAGTGTAGGCAAGATC 840  
DB 781 AATGATGGCTTACTCACACCATTTAAATGTCACTGCCAGCAAGTGTAGGCAAGATC 840  
QY 841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTACTGCA 900  
DB 841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTACTGCA 900  
QY 901 ACATGAGCTGATCTCCATCTTAAAGTAAAGTGTGCTGTACCATTTACACATGCTG 960  
DB 901 ACATGAGCTGATCTCCATCTTAAAGTAAAGTGTGCTGTACCATTTACACATGCTG 960  
QY 961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTCTGTGTAATGC 1020  
DB 961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTCTGTGTAATGC 1020  
QY 1021 AATGGACTTGGCAATTCATCTCTTTCATGAGGAGCTTCTAAGAACAGGGTTGAAGT 1080  
DB 1021 AATGGACTTGGCAATTCATCTCTTTCATGAGGAGCTTCTAAGAACAGGGTTGAAGT 1080  
QY 1081 ATGTTCTCTCTCTTAAAGTTATGCTGCAGACCCCAACACTGCTCAATTTGTACATAAAG 1140  
DB 1081 ATGTTCTCTCTCTTAAAGTTATGCTGCAGACCCCAACACTGCTCAATTTGTACATAAAG 1140  
QY 1141 TGCTTATAGACTTGGCTCCACACACAGTGTAAAGAAAGATTAGCATATGAATTTAAGG 1200  
DB 1141 TGCTTATAGACTTGGCTCCACACACAGTGTAAAGAAAGATTAGCATATGAATTTAAGG 1200  
QY 1201 CCATCTGTTGCTGCAAGCTGCACGAGAGCTGATGTTTACTCGAATCAAAAACATCTCTC 1260  
DB 1201 CCATCTGTTGCTGCAAGCTGCACGAGAGCTGATGTTTACTCGAATCAAAAACATCTCTC 1260  
QY 1261 TCTGGAATGTTGAATTTCAAGCATCTTCAACACATGAAACAGAGATTGTGCTGC 1320  
DB 1261 TCTGGAATGTTGAATTTCAAGCATCTTCAACACATGAAACAGAGATTGTGCTGC 1320  
QY 1321 TGCAATCTCCATATCCCAAGAAAGCAATATGTTGAACCTTGTGCTAAGAAAGAGCAAA 1380  
DB 1321 TGCAATCTCCATATCCCAAGAAAGCAATATGTTGAACCTTGTGCTAAGAAAGAGCAAA 1380  
QY 1381 CATCAATGAAAGACTTAAAGAAATCTTGAACCTCTGCACTGCGCATCTGAGAAAGCTCA 1440  
DB 1381 CATCAATGAAAGACTTAAAGAAATCTTGAACCTCTGCACTGCGCATCTGAGAAAGCTCA 1440  
QY 1441 TAATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGTTAATGCTCTGGATTAAT 1500  
DB 1441 TAATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGTTAATGCTCTGGATTAAT 1500  
QY 1501 TGGTCAGACTTCTTACACAGAGCTGCATATGTTGGTCTATCAAACTGCCGCTACT 1560  
DB 1501 TGGTCAGACTTCTTACACAGAGCTGCATATGTTGGTCTATCAAACTGCCGCTACT 1560  
QY 1561 CCGTAGCTATGGCTGTGATCCCTTAACATTTATATCCCTTCAAGGCTTTACTGCTTTACAGAT 1620  
DB 1561 CCGTAGCTATGGCTGTGATCCCTTAACATTTATATCCCTTCAAGGCTTTACTGCTTTACAGAT 1620  
QY 1621 GGAATGAAATGTACAGCAACTCTCCAGAGGATCTCATTTAGGTAATTTACAGGC 1680  
DB 1621 GGAATGAAATGTACAGCAACTCTCCAGAGGATCTCATTTAGGTAATTTACAGGC 1680





QY	463	ACGTGATGATGGGGCTTATTCCCTTATCATATGATGCTCTTTGGTCATGCTGAAGT	522
Db	482	ACGTGATGATGGGGCTTATTCCCTTATCATATGATGCTCTTTGGTCATGCTGAAGT	541
QY	523	AGTCAATCTCCTTTTTCGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC	582
Db	542	AGTCAATCTCCTTTTTCGACATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTATAC	601
QY	583	TCCTCTCCATGAAGCTGCAATTAAGGAAGATTGATGTTTGCATGCTGCTGTACAGCA	642
Db	602	TCCTCTCCATGAAGCTGCAATTAAGGAAGATTGATGTTTGCATGCTGCTGTACAGCA	661
QY	643	TGGAGCTGAGCCCAACCATCGAATACAGATGGAAGSACAGCATTGGATTTAGCAGATCC	702
Db	662	TGGAGCTGAGCCCAACCATCGAATACAGATGGAAGSACAGCATTGGATTTAGCAGATCC	721
QY	703	ATCTGCCAAGCAGTGTCTTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAGTGCCAG	762
Db	722	ATCTGCCAAGCAGTGTCTTACTGGTGAATATAAGAAAGATGAACCTTTAGAAAGTGCCAG	781
QY	763	GAGTGGCAATGAAGAAATGATGGCTCTACTCACACCATTAATGCTCAACTGCCACGC	822
Db	782	GAGTGGCAATGAAGAAATGATGGCTCTACTCACACCATTAATGCTCAACTGCCACGC	841
QY	823	AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	882
Db	842	AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT	901
QY	883	TGTACAGCTTTTACTGCAACATGGAGCTGATGCTCCATGCTTAAGATAAAGGTGATCTGGT	942
Db	902	TGTACAGCTTTTACTGCAACATGGAGCTGATGCTCCATGCTTAAGATAAAGGTGATCTGGT	961
QY	943	ACCATTACACAATGCCCTGTTCTTATGCTCATATGAAGTAACCTGAACCTTTTGGTCAAGCA	1002
Db	962	ACCATTACACAATGCCCTGTTCTTATGCTCATATGAAGTAACCTGAACCTTTTGGTCAAGCA	1021
QY	1003	TGTTGCTGTGTAATGCAATGGACTTTGTGCAATTCACCTCTTCTCATGAGGAGCTTC	1062
Db	1022	TGTTGCTGTGTAATGCAATGGACTTTGTGCAATTCACCTCTTCTCATGAGGAGCTTC	1081
QY	1063	TAAGAACAGGTTCAAGTATGTTCTTCTTAAAGTTATGCTGCAGACCCCAACACTGCT	1122
Db	1082	TAAGAACAGGTTCAAGTATGTTCTTCTTAAAGTTATGCTGCAGACCCCAACACTGCT	1141
QY	1123	CAATTGTACAAATAAAAGTGTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT	1182
Db	1142	CAATTGTACAAATAAAAGTGTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT	1201
QY	1183	AGCATATGAATTTAAAGGCCACTGTTCTTCAAGCTGCAGAGAAAGCTGATGTTACTCG	1242
Db	1202	AGCATATGAATTTAAAGGCCACTGTTCTTCAAGCTGCAGAGAAAGCTGATGTTACTCG	1261
QY	1243	AATCAAAAACATCTCTCTGGAATTTCAAGCATCTTCAACACATGAAAC	1302
Db	1262	AATCAAAAACATCTCTCTGGAATTTCAAGCATCTTCAACACATGAAAC	1321
QY	1303	AGCATTCATTTGCTGCTGCTCATATCCCAAAAAGAAAGCAATATGTGAAGTGT	1362
Db	1322	AGCATTCATTTGCTGCTGCTCATATCCCAAAAAGAAAGCAATATGTGAAGTGT	1381
QY	1363	GCTAAGAAAAGGAGCAACATCAATGAAGACTTAAAGAAATTTCTTCACTCTCTGACGT	1422
Db	1382	GCTAAGAAAAGGAGCAACATCAATGAAGACTTAAAGAAATTTCTTCACTCTCTGACGT	1441
QY	1423	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAAGCAAGAGT	1482
Db	1442	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAAGCAAGAGT	1501
QY	1483	TAATGCTCTGGATAATCTTGGTTCACATCTCTACACAGAGCTGCATATTTGGTGCATCT	1542
Db	1502	TAATGCTCTGGATAATCTTGGTTCACATCTCTACACAGAGCTGCATATTTGGTGCATCT	1561
QY	1543	ACAAAGCTCCGCCCTTACTCTCTGAGCTATGGGTGTGATCCTTAACATATATATCCCTTCAGGG	1602

RESULT 2  
AAA91488  
ID AAA91488 standard; DNA; 3816 BP.  
XX  
AC AAA91488;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Tankyrase homologue isotype 2 coding sequence.  
XX  
KW Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;  
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;  
KW chromosome mapping; gene therapy; vaccine; ds.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT CDS 2..3726  
FT /\*tag= a  
FT /partial  
FT /product= "TaHo2"  
FT /note= "Tankyrase homologue isotype 2; No start  
FT codon given"  
XX  
PN WO200130987-A2.  
XX  
PD 03-MAY-2001.  
XX  
PF 25-OCT-2000; 2000WO-US41528.  
XX  
PR 25-OCT-1999; 99US-0427154.  
XX  
PA (RIGE-) RIGEL PHARM INC.  
XX  
PI Luo Y, Chan E, Xu X, Huang B;  
XX  
DR WPI: 2001-300503/31.  
DR P-PSDB; AAY97749.  
XX  
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for  
PT inducing or preventing cell proliferation in cells, and for diagnosing,  
PT treating or preventing cell cycle associated disorders such as cancer  
PT  
XX  
PS Claim 13; Fig 2; 63pp; English.  
XX  
CC This sequence encodes the Tankyrase homologue isotype 2 (TaHo-2) protein  
CC of the invention. The invention also relates to the TaHo-1 protein.  
CC The TaHo proteins are useful for inducing or preventing cell  
CC proliferation in cells, and in the study or treatment of conditions  
CC mediated by the cell cycle proteins, such as to diagnose, treat or  
CC prevent cell cycle associated disorders, preferably cancer. The TaHo  
CC coding sequences are useful as hybridisation probes, in chromosome and  
CC gene mapping and in the generation of anti-sense DNA and RNA. The coding  
CC sequences are also useful for the preparation of TaHo, for generating  
CC either transgenic animals or knock out animals which, in turn, are useful  
CC in a development and screening of therapeutically useful agents, in gene  
CC therapy, as vaccine, and for construction of hybridisation probes for  
CC mapping the gene which encodes TaHo and for the genetic analysis of  
CC individuals with genetic disorders. The TaHo proteins, and their coding  
CC sequences are useful in screening assays.  
XX  
SQ Sequence 3816 BP; 1097 A; 806 C; 949 G; 964 T; 0 other;  
Query Match 89.4%; Score 3393.4; DB 22; Length 3816;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 403 AGGTTTGGCGGAAGACGTAGTTGAATATTTGCTTCAGAAATGTCGAAGTGTCCAAAGC 462  
DB 422 AGGTTTGGCGGAAGACGTAGTTGAATATTTGCTTCAGAAATGTCGAAGTGTCCAAAGC 481



Db 1562 ACAAACTGCGGCTACTCTCTGAGCTATGGGTGTGATCTCAATATATATCTCCCTTCAGGG 1621  
QY CTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCTC 1662  
Db 1622 CTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCTC 1681  
QY ATTAGTAAATTCAGAGCGACAGACAAATTCGTTGGAAGCTGCAAGGCTGGAGATGCGA 1722  
Db 1682 ATTAGTAAATTCAGAGCGACAGACAAATTCGTTGGAAGCTGCAAGGCTGGAGATGCGA 1741  
QY AACTGTAAAAAACTGTGTACTGTTCCAGAGTGTCAACTCGAGAGACATTTGAAGGGCGTCA 1782  
Db 1742 AACTGTAAAAAACTGTGTACTGTTCCAGAGTGTCAACTCGAGAGACATTTGAAGGGCGTCA 1801  
QY GTCTACACCACTTCATTTGTCAGCTGGGTATACAGAGTGTCCGTGGTGGTGAATATCTGCT 1842  
Db 1802 GTCTACACCACTTCATTTGTCAGCTGGGTATACAGAGTGTCCGTGGTGGTGAATATCTGCT 1861  
QY ACAGCATGGAGCTGATGTGCTAAAGATAAAGAGGCTTGTACCTTTGCACAAATGC 1902  
Db 1862 ACAGCATGGAGCTGATGTGCTAAAGATAAAGAGGCTTGTACCTTTGCACAAATGC 1921  
QY ATGTTCTTATGGACATATGAAGTTGCAAGCTTCTGTTAAACATGGAAGCTAGTTAA 1962  
Db 1922 ATGTTCTTATGGACATATGAAGTTGCAAGCTTCTGTTAAACATGGAAGCTAGTTAA 1981  
QY TGTAGTGTATTTATGAAATTTACACTTTACATGAAGCAGCAGCAAAAGGAAATATCA 2022  
Db 1982 TGTAGTGTATTTATGAAATTTACACTTTACATGAAGCAGCAGCAAAAGGAAATATCA 2041  
QY AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTACCAAAAAAACAGGATGGAAA 2082  
Db 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTACCAAAAAAACAGGATGGAAA 2101  
QY TACTCTTTGGATCTTTAAAGATGGAGATACAGATATTCAGATCTCTTAGGGGGA 2142  
Db 2102 TACTCTTTGGATCTTTAAAGATGGAGATACAGATATTCAGATCTCTTAGGGGGA 2161  
QY TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTCAAGAACTGCTCTTC 2202  
Db 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTCAAGAACTGCTCTTC 2221  
QY TCTGTATATGTAATTTGCGGATATCCCAAGCAGACATTCACACCTTTACATTTAGC 2262  
Db 2222 TCTGTATATGTAATTTGCGGATATCCCAAGCAGACATTCACACCTTTACATTTAGC 2281  
QY AGCTGTTTATATTAATTTAGAACTTGCAGATATTTGTTTACACACGAGCTGATGTGA 2322  
Db 2282 AGCTGTTTATATTAATTTAGAACTTGCAGATATTTGTTTACACACGAGCTGATGTGA 2341  
QY TGCCCAAGACAAAGGAGGACTTATCTCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382  
Db 2342 TGCCCAAGACAAAGGAGGACTTATCTCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401  
QY TGTAGCAGCTCTACTAATTAAGTATTAATGCTATGCTCAATGCCAGGCAAAATGGGCTTT 2442  
Db 2402 TGTAGCAGCTCTACTAATTAAGTATTAATGCTATGCTCAATGCCAGGCAAAATGGGCTTT 2461  
QY CACACCTTTGCAAGAGCAGCCCAAGAGGAGCAACAGCTTTCTGCTTTGCTGCTAGC 2502  
Db 2462 CACACCTTTGCAAGAGCAGCCCAAGAGGAGCAACAGCTTTCTGCTTTGCTGCTAGC 2521  
QY CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
Db 2522 CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581  
QY AGCGGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCCATGCTGTCGCCCTCTTG 2622  
Db 2582 AGCGGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCCATGCTGTCGCCCTCTTG 2641  
QY TTACAGGCTCAAGTCTCAATGTTGTAGAGGCCAGGAGCCATGCGAGATGCTCTCTC 2682

Db 2642 TTACAAGCCTCAAGTGTCTCAATGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC 2701  
QY TTCAGGTCATCTACCCCATCAAGCCTTCTGACGCCAGCAGCTCTTGACAACTTTATCTGG 2742  
Db 2702 TTCAGGTCATCTACCCCATCAAGCCTTCTGACGCCAGCAGCTCTTGACAACTTTATCTGG 2761  
QY GAGTTCCTTTCAGAACTGCTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTTT 2802  
Db 2762 GAGTTCCTTTCAGAACTGCTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTTT 2821  
QY GGAGAAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAACTCTTGG 2862  
Db 2822 GGAGAAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGAACTCTTGG 2881  
QY ACTTCAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922  
Db 2882 ACTTCAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941  
QY GATGGGCGACAAAGAGCTGAAGGAGTTGGAATCAATCTTTATGGACATAGGCACAACT 2982  
Db 2942 GATGGGCGACAAAGAGCTGAAGGAGTTGGAATCAATCTTTATGGACATAGGCACAACT 3001  
QY AATTAAGAGGTCGAGAGACTTATCTCCGGACAAACAGGCTCTTAACCCATATTTAACTTT 3042  
Db 3002 AATTAAGAGGTCGAGAGACTTATCTCCGGACAAACAGGCTCTTAACCCATATTTAACTTT 3061  
QY GAACACCTCTGCTAGTGAACAATCTTATAGATCTCTCTGATGATATAAGAGTTTCA 3102  
Db 3062 GAACACCTCTGCTAGTGAACAATCTTATAGATCTCTCTGATGATATAAGAGTTTCA 3121  
QY GTCTGTGGAGGAAGAGATGCAAGGTACAGTTCGAGAGCAGAGATGGAGTCTATGAGG 3162  
Db 3122 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTCGAGAGCAGAGATGGAGTCTATGAGG 3181  
QY TGGAACTCTTCAACAGATACAATATCTCAAGATTTCAGAAAGTTTGTAAACAAAGAACTATG 3222  
Db 3182 TGGAACTCTTCAACAGATACAATATCTCAAGATTTCAGAAAGTTTGTAAACAAAGAACTATG 3241  
QY GGAAGATACACTCACCGCAGAAAAGTTCCTGAGAAAAACCAACCAATGCCAATGA 3282  
Db 3242 GGAAGATACACTCACCGCAGAAAAGTTCCTGAGAAAAACCAACCAATGCCAATGA 3301  
QY ACGAATGCTATTTTCATGSGTCTCCTTTTGTGAATGCAATATTCACAAAGGCTTTTGATGA 3342  
Db 3302 ACGAATGCTATTTTCATGSGTCTCCTTTTGTGAATGCAATATTCACAAAGGCTTTTGATGA 3361  
QY AAGCATGCGTACATAGTGTATGTTTGGAGCTGGCAATTTATTTTGGTGAAGAACTCTTC 3402  
Db 3362 AAGCATGCGTACATAGTGTATGTTTGGAGCTGGCAATTTATTTTGGTGAAGAACTCTTC 3421  
QY CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTGTCAGTTTCAACAAGA 3462  
Db 3422 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGGTGTCAGTTTCAACAAGA 3481  
QY CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCCGGGTAACTTTGGGAAAGTC 3522  
Db 3482 CAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCCGGGTAACTTTGGGAAAGTC 3541  
QY TTTCTCTGAGTTCAGTGCAATGAAATGGCACATTCCTCCAGGTCATCACTCAGTCAAC 3582  
Db 3542 TTTCTCTGAGTTCAGTGCAATGAAATGGCACATTCCTCCAGGTCATCACTCAGTCAAC 3601  
QY TGGTAGGCCAGTGTAAATGGCCTAGCATTTAGCTGAATATGTTATTTTACAGAGGAGAA 3642  
Db 3602 TGGTAGGCCAGTGTAAATGGCCTAGCATTTAGCTGAATATGTTATTTTACAGAGGAGAA 3661  
QY GGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTTAGGCCCTGAAGGTATGTCGATGG 3702  
Db 3662 GGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTTAGGCCCTGAAGGTATGTCGATGG 3721  
QY ATAAATAGTTATTTTAAAGAACTAAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGG 3762  
Db 3722 ATAAATAGTTATTTTAAAGAACTAAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGG 3781

QY 3763 CCTCTACGTTTACTCCTTTGCTGAAAAA 3797  
 DB 3782 CCTCTACGTTTACTCCTTTGCTGAAAAA 3816

RESULT 3  
 AAC66825  
 ID AAC66825 standard; cDNA; 4275 BP.  
 XX AAC66825;  
 AC AAC66825;  
 XX  
 XX 27-FEB-2001 (first entry)  
 DT  
 DE Human tankyrase II coding sequence SEQ ID NO: 5.  
 XX Human; tankyrase II; telomere length; signal transduction; ss.  
 KW Human  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 284..3784  
 FT /\*tag= a  
 FT /product= "tankyrase II"  
 FT /transl\_except= (pos:655..660,aa:TyrThr)  
 XX  
 PN WO200061813-A1.  
 XX  
 XX 19-OCT-2000.  
 PD  
 XX  
 PF 10-APR-2000; 2000WO-US09558.  
 XX  
 PR 09-APR-1999; 99US-0128577.  
 PR 13-APR-1999; 99US-0129123.  
 XX  
 XX (GERO-) GERON CORP.  
 PA  
 XX Morin GB, Funk WD, Piatyszek MA;  
 PI WPI; 2000-679503/66.  
 DR P-PSDB; AAB27211.  
 XX  
 XX Novel mammalian Tankyrase II polypeptide and the polynucleotide  
 PT encoding the polypeptide useful for modulating or maintaining telomere  
 PT length, replicative capacity, apoptosis, chromosome packing or gene  
 PT expression -  
 XX  
 XX Claim 1; Fig 4; 52pp; English.  
 PS  
 XX The present sequence is a version of the human tankyrase II coding  
 CC sequence. Its protein is thought to be involved in signal transduction in  
 CC the cell, and to have binding activity for other telomere-associated  
 CC proteins. It is possible that it plays a role in the regulation of  
 CC telomere length, thus affecting the replicative ability of the cell. The  
 CC protein is useful for ribosylating target proteins, for determining  
 CC tankyrase II binding activity in a sample, and for modulating telomere  
 CC length in a cell.  
 XX  
 SQ Sequence 4275 BP; 1247 A; 889 C; 1025 G; 1114 T; 0 other;

Query Match 89.2%; Score 3388.8; DB 21; Length 4275;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 403 AGCTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCCAAGC 462  
 DB 481 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGTCCCAAGC 540  
 QY 463 ACCTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTGGTGCATGCTGAAGT 522  
 DB 541 ACCTGATGATGGGGCTTATTCCTCTTCATATGATGCTCTTTGGTGCATGCTGAAGT 600  
 QY 523 AGTCAATCTCTTTTCGACATGGTGCAGACCCCAATGCTGCGAGATAATTTGGAATTATAC 582

DB 601 AGTCAATCTCTTTTTCGACATGGTGCAGACCCCAATGCTCGAGATAAATGGAATTATAC 660  
 QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTCATGTTTTCATTTGCTGTACAGCA 642  
 DB 661 TCCTCTCCATGAAGCTGCAATTAAGGAAAGATTCATGTTTTCATTTGCTGTACAGCA 720  
 QY 643 TGGAGCTGAGCCAAACATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 702  
 DB 721 TGGAGCTGAGCCAAACATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 780  
 QY 703 ATCTGCCAAAGCAGTCTTACTTGGTGAATATTAAGAAAGATGAACCTCTTGAAGTCCAG 762  
 DB 781 ATCTGCCAAAGCAGTCTTACTGGTGAATATTAAGAAAGATGAACCTCTTGAAGTCCAG 840  
 QY 763 GAGTGGCAATGAAGAAAAAATGATGGCTTACTTACACCACTTAAATGTCAACTGCCACGC 822  
 DB 841 GAGTGGCAATGAAGAAAAAATGATGGCTTACTTACACCACTTAAATGTCAACTGCCACGC 900  
 QY 823 AAGTATGCGAGAAAGTCAACTCCATTAATTTGGCAGCAGGATATAACAGAGTAAAGAT 882  
 DB 901 AAGTATGCGAGAAAGTCAACTCCATTAATTTGGCAGCAGGATATAACAGAGTAAAGAT 960  
 QY 883 TGTACAGCTGTTTACTGCAACATGGAGCTGATGTCATGCTTAAAGATAAAGGTGATCTGGT 942  
 DB 961 TGTACAGCTGTTTACTGCAACATGGAGCTGATGTCATGCTTAAAGATAAAGGTGATCTGGT 1020  
 QY 943 ACCATTACACAATGCTGTTTATGTCATTAATGAAGTAACCTGAACCTTTTGGTCAAGCA 1002  
 DB 1021 ACCATTACACAATGCTGTTTATGTCATTAATGAAGTAACCTGAACCTTTTGGTCAAGCA 1080  
 QY 1003 TGGTGCCTGTGTAATGCAATGGAGCTTGGCAATTCACCTCTTCATGAGCAGCCTTC 1062  
 DB 1081 TGGTGCCTGTGTAATGCAATGGAGCTTGGCAATTCACCTCTTCATGAGCAGCCTTC 1140  
 QY 1063 TAAGAACAGGTTGAAGTATGTTCTCTCTTCTTAAGTTATGTTGTCAGACCCCAACTGCT 1122  
 DB 1141 TAAGAACAGGTTGAAGTATGTTCTCTCTTCTTAAGTTATGTTGTCAGACCCCAACTGCT 1200  
 QY 1123 CAATTGTCAATTAAGTGTCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1182  
 DB 1201 CAATTGTCAATTAAGTGTCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1260  
 QY 1183 AGCATATGAATTAAGGCCACCTGTTGCTGCAAGCTGCACGAGAGAGCTGATGTTACTCG 1242  
 DB 1261 AGCATATGAATTAAGGCCACCTGTTGCTGCAAGCTGCACGAGAGAGCTGATGTTACTCG 1320  
 QY 1243 AATCAAAAAACATCTCTCTGGAATTTGTTGAAGTAGTGGTGAACATCCTCAAAACATGAAC 1302  
 DB 1321 AATCAAAAAACATCTCTCTGGAATTTGTTGAAGTAGTGGTGAACATCCTCAAAACATGAAC 1380  
 QY 1303 AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTGAATGTT 1362  
 DB 1381 AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTGAATGTT 1440  
 QY 1363 GCTAAGAAAGAGAGCAACATCAATGAAAGACTTAAGAAATTTCTGACTCCTCTGCACGT 1422  
 DB 1441 GCTAAGAAAGAGAGCAACATCAATGAAAGACTTAAGAAATTTCTGACTCCTCTGCACGT 1500  
 QY 1423 GGCACTGTGAGAAAGCTCATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482  
 DB 1501 GGCACTGTGAGAAAGCTCATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAGGT 1560  
 QY 1483 TAATGCTCTGGATAATCTTGGTGCAGACTTCTTACACAGAGCTGCATATTTGGTCACT 1542  
 DB 1561 TAATGCTCTGGATAATCTTGGTGCAGACTTCTTACACAGAGCTGCATATTTGGTCACT 1620  
 QY 1543 ACAAACTGCCGCTACTCTGAGCTATGGTGTGATGCTTAACATATATATCCCTTCAGGG 1602  
 DB 1621 ACAAACTGCCGCTACTCTGAGCTATGGTGTGATGCTTAACATATATATCCCTTCAGGG 1680  
 QY 1603 CTTTACTGCTTTACAGATGGAAATGAAAATGTACAGCAACTCCTCCAAAGAGGTATCTC 1662





Db	1409	GCTAAGAAAAGGAGCAAAACATCAATGAAAGACGTAAAGAAATCTTGACATCCCTCTCGAGGT	1468
Qy	1423	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGGT	1482
Db	1469	GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACACATGAAGCAAGGT	1528
Qy	1483	TAATGCTCTGGATAATCTTTGGTCAGACTCTCTACACAGAGCTGCATATTTGTGGTCATCT	1542
Db	1529	TNATGCTCTGGATAATCTTTGGTCAGACTCTCTACACAGAGCTGCATATTTGTGGTCATCT	1588
Qy	1543	ACAAACCTGCCGCTACTCTGAGCTATGGGTGTGATCCTTAACATATATATPCCCTTCAGGG	1602
Db	1589	ACAAACCTGCCGCTACTCTGAGCTATGGGTGTGATCCTTAACATATATATPCCCTTCAGGG	1648
Qy	1603	CTTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCCAGAGGGTATCTC	1662
Db	1649	CTTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCCAGAGGGTATCTC	1708
Qy	1663	ATTAGGTAAATTCAGAGCGACAGACAATTCGTGGAAGCTGCAAGGCTGGAGATGTCGA	1722
Db	1709	ATTAGGTAAATTCAGAGCGACAGACAATTCGTGGAAGCTGCAAGGCTGGAGATGTCGA	1768
Qy	1723	AACTGTAAAAAACTGTGTACTTGTTCAGAGTGTCAACTGCAGAGACAATGAAGGGCGTCA	1782
Db	1769	AACTGTAAAAAACTGTGTACTTGTTCAGAGTGTCAACTGCAGAGACAATGAAGGGCGTCA	1828
Qy	1783	GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT	1842
Db	1829	GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT	1888
Qy	1843	ACAGCATGGAGCTGATGTCGATCTGCTAAAGATAAAGAGGGCCCTGTACCTTTGTCACAAATGC	1902
Db	1889	ACACCATGGAGCTGATGTCGATCTGCTAAAGATAAAGAGGGCCCTGTACCTTTGTCACAAATGC	1948
Qy	1903	ATGTTCTTATGGACATATTAAGATTGCGAACAATCTTGTGTTAAACATGGAGCACTAGTTAA	1962
Db	1949	ATGTTCTTATGGACATATTAAGATTGCGAACAATCTTGTGTTAAACATGGAGCACTAGTTAA	2008
Qy	1963	TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAACAAAGAAATATGA	2022
Db	2009	TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAACAAAGAAATATGA	2068
Qy	2023	AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAGGATGGAAA	2082
Db	2069	AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAGGATGGAAA	2128
Qy	2083	TACTTCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAGAAATCTGCTTAGGGGAGA	2142
Db	2129	TACTTCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTCAGAAATCTGCTTAGGGGAGA	2188
Qy	2143	TGCAGCTTGTCTAGATGCTGCCAAGAAGGGTTGTTTACCCAGAGTGAAGAAGTTGCTTTC	2202
Db	2189	TGCAGCTTGTCTAGATGCTGCCAAGAAGGGTTGTTTACCCAGAGTGAAGAAGTTGCTTTC	2248
Qy	2203	TCCTGATAATGTAAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAGC	2262
Db	2249	TCCTGATAATGTAAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAGC	2308
Qy	2263	AGCTGGTTATAATAATTTAGAGTTTCGAGATATTTGTTTCAACACAGGAGCTGATGTGAA	2322
Db	2309	AGCTGGTTATAATAATTTAGAGTTTCGAGATATTTGTTTCAACACAGGAGCTGATGTGAA	2368
Qy	2323	TGCCCCAAGACAAGGAGCTTATTCCTTTACATATGCACGATCTTACGGCATGTGAGA	2382
Db	2369	TGCCCCAAGACAAGGAGCTTATTCCTTTACATATGCACGATCTTACGGCATGTGAGA	2428
Qy	2383	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACGACAAATGGGCTTT	2442
Db	2429	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACGACAAATGGGCTTT	2488
Qy	2443	CACACCTTTGCAGGAAGCAGCCCCAAAGGGAGCAACAGACTTTGTGCTTTGTTGCTTAGC	2502



|||||  
Db 1348 ACAAACTCGCGCTACTCTCTGAGTATGGGTGATCCTTAACATTATATCTCCCTTCAGG 1407  
QY 1603 CTTTACTCTCTTACAGATGGGAATGAAATGTACAGCACTCTCTCAAGAGGTAFTCTC 1662  
Db 1408 CTTTACTCTCTTACAGATGGGAATGAAATGTACAGCACTCTCTCAAGAGGTAFTCTC 1467  
QY 1663 ATTAGGTAATTCAGAGGCAGACAGCAATTCGTGGAAGCTGCAAGGCTGGAGATGTCGA 1722  
Db 1468 ATTAGGTAATTCAGAGGCAGACAGCAATTCGTGGAAGCTGCAAGGCTGGAGATGTCGA 1527  
QY 1723 AACTGTAAAAAATCTGTACTGTTACAGATGTCACATGCGAGACATTCGAAGGGGCTCA 1782  
Db 1528 AACTGTAAAAAATCTGTACTGTTACAGATGTCACATGCGAGACATTCGAAGGGGCTCA 1587  
QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATACAGAGTCTCGGTGGTGAATATCTGCT 1842  
Db 1588 GTCTACACCACTTCATTTTGCAGCTGGGTATACAGAGTCTCGGTGGTGAATATCTGCT 1647  
QY 1843 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTGTACCTTTGCGACAAATGC 1902  
Db 1648 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTGTACCTTTGCGACAAATGC 1707  
QY 1903 ATGTTCTTATGGACATTCGAAGTTCAGAACTTCCTGTTAAACATGCGAGCACTAGTTAA 1962  
Db 1708 ATGTTCTTATGGACATTCGAAGTTCAGAACTTCCTGTTAAACATGCGAGCACTAGTTAA 1767  
QY 1963 TGTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCAGCAAAAGGAATATGA 2022  
Db 1768 TGTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCAGCAAAAGGAATATGA 1827  
QY 2023 AATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAACAAGGATGCAAA 2082  
Db 1828 AATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAACAAGGATGCAAA 1887  
QY 2083 TACTCCTTTGGATCTGTTAAAGATGGAGTACAGATATTCAGATCTGCTTAGGGGAGA 2142  
Db 1888 TACTCCTTTGGATCTGTTAAAGATGGAGTACAGATATTCAGATCTGCTTAGGGGAGA 1947  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTCAAGAAGTTGCTTTC 2202  
Db 1948 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTCAAGAAGTTGCTTTC 2007  
QY 2203 TCCTGTAATATGTAATTCGCCGCATACCAAGCAGACATTCACACCTTTACATTTAGC 2262  
Db 2008 TCCTGTAATATGTAATTCGCCGCATACCAAGCAGACATTCACACCTTTACATTTAGC 2067  
QY 2263 AGCTGGTTATAATAATTTAGAACTTCAGAGTATTTGTTACACACGGAGCTGATGTGAA 2322  
Db 2068 AGCTGGTTATAATAATTTAGAACTTCAGAGTATTTGTTACACACGGAGCTGATGTGAA 2127  
QY 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382  
Db 2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187  
QY 2383 TGTAGCAGCTCTACTAATAAAGTATAATGATTAATGCTGTGTCATGCGGACAAATGGGCTTT 2442  
Db 2188 TGTAGCAGCTCTACTAATAAAGTATAATGATTAATGCTGTGTCATGCGGACAAATGGGCTTT 2247  
QY 2443 CACACCTTTGCAAGAGCAGCCCAAGAGGAGCAACACAGCTTGTGCTTCTGCTAGC 2502  
Db 2248 CACACCTTTGCAAGAGCAGCCCAAGAGGAGCAACACAGCTTGTGCTTCTGCTAGC 2307  
QY 2503 CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
Db 2308 CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367  
QY 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTG 2622  
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTG 2427  
QY 2623 TTAAGAGCCTCAAGTCTCAATGGTGTGAGAAGCCAGAGCCACTGCAGATGCTCTCTC 2682  
|||||

Db 2428 TTACAAAGCCTCAAGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTC 2487  
QY 2683 TTCAGTTCATCTAGCCCATCAAGCCTTTTCTGCAGCCAGCAGTCTTTGACAACATTAFTCTCG 2742  
Db 2488 TTCAGTTCATCTAGCCCATCAAGCCTTTTCTGCAGCCAGCAGTCTTTGACAACATTAFTCTCG 2547  
QY 2743 GAGTTTTTTCAGAACTCTCTCTCAGTAGTTAGTTCAAGTGGACAGAGGGTCTTCCAGTTT 2802  
Db 2548 GAGTTTTTTCAGAACTCTCTCTCAGTAGTTAGTTCAAGTGGACAGAGGGTCTTCCAGTTT 2607  
QY 2803 GGAGAAAAAGGAGTTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAAGAACTCTCG 2862  
Db 2608 GGAGAAAAAGGAGTTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAAGAACTCTCG 2667  
QY 2863 ACTTGAGCACTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922  
Db 2668 ACTTGAGCACTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727  
QY 2923 GATGGGCAACAAGGAGCTCAAGGAGATTGGAATCAATGCTTATGACATAGGCACAAACT 2982  
Db 2728 GATGGGCAACAAGGAGCTCAAGGAGATTGGAATCAATGCTTATGACATAGGCACAAACT 2787  
QY 2983 AATTAAGGAGTCGAGAGACTTATCTCCGCAACAAGGCTTTAACCCATATTTAACTTT 3042  
Db 2788 AATTAAGGAGTCGAGAGACTTATCTCCGCAACAAGGCTTTAAACCCATATTTAACTTT 2847  
QY 3043 GAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATAAGAGTTTCA 3102  
Db 2848 GAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATAAGAGTTTCA 2907  
QY 3103 GTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCGAGAGCACAGAGATGAGGTCATGCAGG 3162  
Db 2908 GTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCGAGAGCACAGAGATGAGGTCATGCAGG 2967  
QY 3163 TGGAACTTTCACAGATCAATATTTCTCAAGATTTCAGAAAGTTTGTAAACAAGAACTATG 3222  
Db 2968 TGGAACTTTCACAGATCAATATTTCTCAAGATTTCAGAAAGTTTGTAAACAAGAACTATG 3027  
QY 3223 GGAAAGATACACTCCCGGAGAAAAGAGTTTCTCAAGAAAACACACCATGCCAATGA 3282  
Db 3028 GGAAAGATACACTCCCGGAGAAAAGAGTTTCTCAAGAAAACACACCATGCCAATGA 3087  
QY 3283 ACGAATGCTATTTCATGGGTCTCTCTTTTGTGAATCAATATTCACAAAGCTTTTGATGA 3342  
Db 3088 ACGAATGCTATTTCATGGGTCTCTCTTTTGTGAATCAATATTCACAAAGCTTTTGATGA 3147  
QY 3343 AAGCATGCGTACATPAGTGTGTATGTTTGGAGCTGGCAATTTATTTGCTGAAAACCTCTTC 3402  
Db 3148 AAGCATGCGTACATPAGTGTGTATGTTTGGAGCTGGCAATTTATTTGCTGAAAACCTCTTC 3207  
QY 3403 CAAAAGCAATCAATATGTATGGAATTTGGAGGAGGTAAGTGGGTCCAGTTCCAAAGA 3462  
Db 3208 CAAAAGCAATCAATATGTATGGAATTTGGAAATTTGGAGGAGGTAAGTGGGTCCAGTTCCAAAGA 3267  
QY 3463 CAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 3522  
Db 3268 CAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 3327  
QY 3523 TTTCTGTCAGTTCAGTGCATGCAATGAAATGGCACATTCCTCCAGGTCACTACTCAGTCAC 3582  
Db 3328 TTTCTGTCAGTTCAGTGCATGCAATGAAATGGCACATTCCTCCAGGTCACTACTCAGTCAC 3387  
QY 3583 TGGTAGGCCCCAGTGAATGCTTAGCATTAGCTGAATATGTTATTTACAGAGAGAGAACA 3642  
Db 3388 TGGTAGGCCCCAGTGAATGCTTAGCATTAGCTGAATATGTTATTTACAGAGAGAGAACA 3447  
QY 3643 GGCTTATCTCTGAGTATTTAATTACTTACCGAGTTATCAGGCTCAAGGTATGTCGATGG 3702  
Db 3448 GGCTTATCTCTGAGTATTTAATTACTTACCAGATTATCAGGCTCAAGGTATGTCGATGG 3507  
QY 3703 ATAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAATCAATCAAAAGCAGCTGG 3762  
Db 3508 ATAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAATCAATCAAAAGCAGCTGG 3567



QY 3763 CCTCTACGTTTACTCCTTTGCTGAAAAAAA 3794  
 DB 3568 CCTCTACGTTTACTCCTTTGCTGAAAAAAA 3599

RESULT 6  
 AAF63930  
 ID AAF63930 standard; DNA; 4406 BP.  
 AC AAF63930;  
 XX  
 XX 05-APR-2001 (first entry)  
 XX  
 DE Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 106.  
 XX  
 KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
 KW inflammatory disorder; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200100849-A1.  
 XX  
 XX 04-JAN-2001.  
 XX  
 XX 28-JUN-2000; 2000WO-US17827.  
 XX  
 XX 29-JUN-1999; 99US-0141582.  
 XX  
 XX (ICOS-) ICOS CORP.  
 XX  
 XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 PI  
 XX WPI; 2001-102896/11.  
 DR  
 DR P-PSDB; AAB66290.

XX  
 XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
 PT inflammatory and autoimmune disorders -  
 XX  
 XX Example 2; Page 168-173; 242pp; English.

XX  
 XX The present invention provides the protein and coding sequence for the  
 CC human tankyrase2 protein. This is found in two different versions,  
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
 CC polyADP-ribosylation activity and is involved in the modification of  
 CC TRF1, which is a telomere-specific binding protein. The regulation of  
 CC telomere length, in which TRF1 has a role, is linked to ageing and  
 CC cancer. The sequences are useful in the treatment of cancers and  
 CC inflammatory disorders.

XX  
 XX Sequence 4406 BP; 1268 A; 910 C; 1054 G; 1174 T; 0 other;  
 SQ

Query Match 89.2%; Score 3387.2; DB 22; Length 4406;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 403 AGGTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGCCAAAGC 462  
 DB 487 AGGTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAGAAATGGTGCAAGTGCCAAAGC 546  
 QY 463 ACGTGATGATGGGGCCCTTATTCCTCTTCATTAATGCAATGCTTTTGGTCATGCTGAAGT 522  
 DB 547 ACGTGATGATGGGGCCCTTATTCCTCTTCATTAATGCAATGCTTTTGGTCATGCTGAAGT 606  
 QY 523 AGTCAATCTCCTTTGGACATGCTGCAGACCCCAATGCTCGAGATTAATGGAATTATAC 582  
 DB 607 AGTCAATCTCCTTTGGACATGCTGCAGACCCCAATGCTCGAGATTAATGGAATTATAC 666  
 QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAAGCAATGATGTTTGGCATTTGCTGTTTACAGCA 642  
 DB 667 TCCTCTCCATGAAGCTGCAATTAAGGAAAGCAATGATGTTTGGCATTTGCTGTTTACAGCA 726

QY 643 TGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGACAGCAGCATTTGATTTAGCAGATCC 702  
 DB 727 TGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGACAGCAGCATTTGATTTAGCAGATCC 786  
 QY 703 ATCTGCCAAAGCAGTCTTACTGGTGAATATAGAAGATGAACCTCTTAGAAGAGTCCAG 762  
 DB 787 ATCTGCCAAAGCAGTCTTACTGGTGAATATAGAAGATGAACCTCTTAGAAGAGTCCAG 846  
 QY 763 GAGTGGCAATGAAGAAAAATGATGCTCTACTCACACCATTAATGTCAACTGCCACGC 822  
 DB 847 GAGTGGCAATGAAGAAAAATGATGCTCTACTCACACCATTAATGTCAACTGCCACGC 906  
 QY 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGAT 882  
 DB 907 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGAT 966  
 QY 883 TGTACAGCTTTACTGCAACATGGAGCTGATGTCCTTAAGTAAAGATAAAGGTGATCTGGT 942  
 DB 967 TGTACAGCTTTACTGCAACATGGAGCTGATGTCCTTAAGTAAAGATAAAGGTGATCTGGT 1026  
 QY 943 ACCATTACACAATGCCTGTTCTTATGCTCATTAAGTAACCTGAACCTTTGGTCAAGCA 1002  
 DB 1027 ACCATTACACAATGCCTGTTCTTATGCTCATTAAGTAACCTGAACCTTTGGTCAAGCA 1086  
 QY 1003 TGTGCTCTGTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1062  
 DB 1087 TGTGCTCTGTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1146  
 QY 1063 TAAGAACAGGGTTGAAGTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1122  
 DB 1147 TAAGAACAGGGTTGAAGTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1206  
 QY 1123 CAATTGTCACAATAAAGTGTCTATAGACTTTGGCTCCACACCACACAGTTAAAGAAAGATT 1182  
 DB 1207 CAATTGTCACAATAAAGTGTCTATAGACTTTGGCTCCACACCACACAGTTAAAGAAAGATT 1266  
 QY 1183 AGCATATGAATTAAGGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1242  
 DB 1267 AGCATATGAATTAAGGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326  
 QY 1243 AATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCTCAACACATGAAC 1302  
 DB 1327 AATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCTCAACACATGAAC 1386  
 QY 1303 AGCATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1362  
 DB 1387 AGCATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446  
 QY 1363 GCTAAGAAAAGGACCAACATCAATGAAGACCTAAGAAATTTCTTGACTCTCTCTGCACTG 1422  
 DB 1447 GCTAAGAAAAGGACCAACATCAATGAAGACCTAAGAAATTTCTTGACTCTCTCTGCACTG 1506  
 QY 1423 GGCATCTGAGAAAGCTCATATATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482  
 DB 1507 GGCATCTGAGAAAGCTCATATATGATTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1566  
 QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGCTGCTCATCT 1542  
 DB 1567 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGCTGCTCATCT 1626  
 QY 1543 ACRAACCTGCCGCCCTACTCTGAGCTATGGGTGATGCTTAACTATATATCCCTTCAGGG 1602  
 DB 1627 ACRAACCTGCCGCCCTACTCTGAGCTATGGGTGATGCTTAACTATATATCCCTTCAGGG 1686  
 QY 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCTCTCCAGAGGTATCTC 1662  
 DB 1687 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCTCTCCAGAGGTATCTC 1746  
 QY 1663 ATTAGGTAATTCAGAGCAGACAGCAATTTGCTGGAAGCTCAAGAGCTGGAGATGTGCA 1722  
 DB 1747 ATTAGGTAATTCAGAGCAGACAGCAATTTGCTGGAAGCTCAAGAGCTGGAGATGTGCA 1806  
 QY 1723 AACTGTAAAAAACTGTGTACTGTTTACAGATGTTCAACTGCAGACAGATTTGAAGGGCGTCA 1782

|||||  
Db 1807 AACTGTAAAAAACTGTGTACTGTTCCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1866  
QY 1783 GTCTACACCACTTCATTTGCGACCTGGGTATACACAGTGTGCGTGGTGAATATCTGCT 1842  
Db 1867 GTCTACACCACTTCATTTGCGACCTGGGTATACACAGTGTGCGTGGTGAATATCTGCT 1926  
QY 1843 ACAGCATGGAGCTGTGTGTCATCTAAAGATAAAGAGAGCCCTGTACCTTTGCACAAATGC 1902  
Db 1927 ACAGCATGGAGCTGTGTGTCATCTAAAGATAAAGAGAGCCCTGTACCTTTGCACAAATGC 1986  
QY 1903 ATGTTCTTATGGACATTAAGTGTGCAAGCTTCTTTGTTAAACATGGAGCAGTAGTTAA 1962  
Db 1987 ATGTTCTTATGGACATTAAGTGTGCAAGCTTCTTTGTTAAACATGGAGCAGTAGTTAA 2046  
QY 1963 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA 2022  
Db 2047 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGA 2106  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAGGATGGAAA 2082  
Db 2107 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAGGATGGAAA 2166  
QY 2083 TACTCTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA 2142  
Db 2167 TACTCTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAGA 2226  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGATGCAAGAGTGTCTTC 2202  
Db 2227 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGATGCAAGAGTGTCTTC 2286  
QY 2203 TCTGTAATATGTAATTTGCCGGATACCCAAAGCGACACATTCACACACCTTTACATTTAGC 2262  
Db 2287 TCTGTAATATGTAATTTGCCGGATACCCAAAGCGACACATTCACACACCTTTACATTTAGC 2346  
QY 2263 AGCTGTTTATAATTTAGAAGTTGCAGAGTATTTGTTTACACACGAGGAGCTGATGTGAA 2322  
Db 2347 AGCTGTTTATAATTTAGAAGTTGCAGAGTATTTGTTTACACACGAGGAGCTGATGTGAA 2406  
QY 2323 TGCCCAAGCAAAAGGAGGACTTATCTCTTTACATATGTCAGCATCTTACGGGCATGTAGA 2382  
Db 2407 TGCCCAAGCAAAAGGAGGACTTATCTCTTTACATATGTCAGCATCTTACGGGCATGTAGA 2466  
QY 2383 TGTAGCAGCTCTACTAATAAGTATAATGCATGTGTCAATGCCAGCAAAATGGGCTTT 2442  
Db 2467 TGTAGCAGCTCTACTAATAAGTATAATGCATGTGTCAATGCCAGCAAAATGGGCTTT 2526  
QY 2443 CACACCTTTGCAGAGCAGCCCAAAAGGAGCAACAGCTTTGTGCTTTGCTAGC 2502  
Db 2527 CACACCTTTGCAGAGCAGCCCAAAAGGAGCAACAGCTTTGTGCTTTGCTAGC 2586  
QY 2503 CCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
Db 2587 CCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2646  
QY 2563 AGCGGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTG 2622  
Db 2647 AGCAGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCCCTCTTG 2706  
QY 2623 TTACAAGCCTCAAGTGTCTCAATGTTGTGAGAGCCGAGGAGCCATGCGATGCTCTCTC 2682  
Db 2707 TTACAAGCCTCAAGTGTCTCAATGTTGTGAGAGCCGAGGAGCCATGCGATGCTCTCTC 2766  
QY 2683 TTCAGTCTCATCTAGCCCATCAAGCTTTCTGTCAGCCAGCAGCTTTGACACTTATCTGG 2742  
Db 2767 TTCAGTCTCATCTAGCCCATCAAGCTTTCTGTCAGCCAGCAGCTTTGACACTTATCTGG 2826  
QY 2743 GAGTTTTCAGAACTGTCTTCACTAGTATGTTCAAGTGAAGCAGAGGCTGCTTCCAGTTT 2802  
Db 2827 GAGTTTTCAGAACTGTCTTCACTAGTATGTTCAAGTGAAGCAGAGGCTGCTTCCAGTTT 2886  
QY 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGTTTATAGCATAACTCAATTCGTAAGGAATCTTGG 2862  
|||||

Db 2887 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAAGAACTCTGG 2946  
QY 2863 ACTTGAGCACCTAATGATATATTTTGAGAGAGAAACAGATCACTTTTGGATGATTTAGTTGA 2922  
Db 2947 ACTTGAGCACCTAATGATATATTTTGAGAGAGAAACAGATCACTTTTGGATGATTTAGTTGA 3006  
QY 2923 GATGGGACACAAAGGAGCTGAAGGAGATGGAATCAATGCTTTATGGACATAGGCACAACT 2982  
Db 3007 GATGGGACACAAAGGAGCTGAAGGAGATGGAATCAATGCTTTATGGACATAGGCACAACT 3066  
QY 2983 AATTAAGGAGCTCGAGAGACTTATCTCCGGACAAACAGGCTCTTAACCCATATTTAACTTT 3042  
Db 3067 AATTAAGGAGCTCGAGAGACTTATCTCCGGACAAACAGGCTCTTAACCCATATTTAACTTT 3126  
QY 3043 GAACACCTCTGGTGTGAGCAAAATTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTCA 3102  
Db 3127 GAACACCTCTGGTGTGAGCAAAATTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTCA 3186  
QY 3103 GTCTGTGGAGAAAGAGATGCAAGGTACAGTTTCAGAGTTCAGAGGCTTTCTAACAAGAACTATG 3162  
Db 3187 GTCTGTGGAGAAAGAGATGCAAGGTACAGTTTCAGAGTTCAGAGGCTTTCTGAGGCTATGAGG 3246  
QY 3163 TGGAACTCTTCAACAGATACAATATTTCTCAAGATTCAGAGGCTTTCTAACAAGAACTATG 3222  
Db 3247 TGGAACTCTTCAACAGATACAATATTTCTCAAGATTCAGAGGCTTTCTAACAAGAACTATG 3306  
QY 3223 GGAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAACCAACACCATGCAATGA 3282  
Db 3307 GGAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAACCAACACCATGCAATGA 3366  
QY 3283 ACGAATGCTATTTTCATGGTCTCTCTTTGCAATGCAATTCAGAGTTCAGAGGCTTTCTAACAAGAACTATG 3342  
Db 3367 ACGAATGCTATTTTCATGGTCTCTCTTTGCAATGCAATTCAGAGTTCAGAGGCTTTCTAACAAGAACTATG 3426  
QY 3343 AAGCATCGGTACATGAGTGTGATGTTTGGAGCTGGCATTTATTTGCTGAAAACTCTTC 3402  
Db 3427 AAGCATCGGTACATGAGTGTGATGTTTGGAGCTGGCATTTATTTGCTGAAAACTCTTC 3486  
QY 3403 CAAGAAGCAATCAATATGATATGAAATTTGGAGGAGTACTGGGTGTCAGTTTCAACAAGA 3462  
Db 3487 CAAGAAGCAATCAATATGATATGAAATTTGGAGGAGTACTGGGTGTCAGTTTCAACAAGA 3546  
QY 3463 CAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 3522  
Db 3547 CAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 3606  
QY 3523 TTTCTGTCAGTTCAGTGCATGCAATGAAATGGCACAATCTCTCCAGGTCACTCAGTCAC 3582  
Db 3607 TTTCTGTCAGTTCAGTGCATGCAATGAAATGGCACAATCTCTCCAGGTCACTCAGTCAC 3666  
QY 3583 TGGTAGGCCCACTGTAATGGCCTAGCATTAGCTGAATATGTTTATTTACAGAGGAGAA 3642  
Db 3667 TGGTAGGCCCACTGTAATGGCCTAGCATTAGCTGAATATGTTTATTTACAGAGGAGAA 3726  
QY 3643 GGCTTATCTCTGAGTATTTAAATTTACTTACCGATTATGAGGCTCAAGGTATGGTTCGATGG 3702  
Db 3727 GGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTATGAGGCTCAAGGTATGGTTCGATGG 3786  
QY 3703 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGAGCAGCAGTGG 3762  
Db 3787 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGAGCAGCAGTGG 3846  
QY 3763 CCTCTAGCTTTTACTCTCTTCTGCTGAAAAAAA 3794  
Db 3847 CCTCTAGCTTTTACTCTCTTCTGCTGAAAAAAA 3878

RESULT 7

AAF63953

ID AAF63953 standard; DNA: 4992 BP.

XX

AC AAF63953;

XX



|||||  
Db 2573 ATGTTCTTATGGACATTTATGAAGTTGCGAAGCTTCTGTTTAAACATGGACGAGTGTAA 2632  
|||||  
QY 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAGCAGCAGCAAAAGAAATATGA 2022  
|||||  
Db 2633 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAGCAGCAGCAAAAGAAATATGA 2692  
|||||  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACCAAAAAACAGGATGAAA 2082  
|||||  
Db 2693 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACCAAAAAACAGGATGAAA 2752  
|||||  
QY 2083 TACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142  
|||||  
Db 2753 TACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2812  
|||||  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGATGCAAGAAGTTGCTTC 2202  
|||||  
Db 2813 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGATGCAAGAAGTTGCTTC 2872  
|||||  
QY 2203 TCTGTATAATGTAATTTGCCCGATACCCAAAGCAGACATTTCAACACCTTTTACATTTAGC 2262  
|||||  
Db 2873 TCTGTATAATGTAATTTGCCCGATACCCAAAGCAGACATTTCAACACCTTTTACATTTAGC 2932  
|||||  
QY 2263 AGCTGGTTATAATTTAGAACTTGCAGAGTATTTGTTTACACACGGAGCTGATGTGAA 2322  
|||||  
Db 2933 AGCTGGTTATAATTTAGAACTTGCAGAGTATTTGTTTACACACGGAGCTGATGTGAA 2992  
|||||  
QY 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTTACATAATGAGATCTTACGGCATGTAGA 2382  
|||||  
Db 2993 TGCCCAAGACAAAGGAGGACTTATTCCTTTTACATAATGAGATCTTACGGCATGTAGA 3052  
|||||  
QY 2383 TGTAGCAGCTCTACTAATAAAGTATAATGATGTGCAATGCCACGACAAATGGGCTTT 2442  
|||||  
Db 3053 TGTAGCAGCTCTACTAATAAAGTATAATGATGTGCAATGCCACGACAAATGGGCTTT 3112  
|||||  
QY 2443 CACACCTTTGCAGAGAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTGTAGC 2502  
|||||  
Db 3113 CACACCTTTGCAGAGAGCAGCCCAAAAGGAGCAACACAGCTTTTGTGCTGTAGC 3172  
|||||  
QY 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGCAGACAAACACCTTTAGATTTAGTTTC 2562  
|||||  
Db 3173 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGCAGACAAACACCTTTAGATTTAGTTTC 3232  
|||||  
QY 2563 AGCGATGATGTGAGCGCTCTCTGACAGAGCCATGCCCCCATCTGCTCCCTCTTG 2622  
|||||  
Db 3233 AGCAGATGATGTGAGCGCTCTCTGACAGAGCCATGCCCCCATCTGCTCCCTCTTG 3292  
|||||  
QY 2623 TTACAAGCTTCAAGTCTCAATGCTGTGAGAGCCAGAGCCACTGCAGATGCTCTCTC 2682  
|||||  
Db 3293 TTACAAGCTTCAAGTCTCAATGCTGTGAGAGCCAGAGCCACTGCAGATGCTCTCTC 3352  
|||||  
QY 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGCTTTGACAACCTTATCTGG 2742  
|||||  
Db 3353 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGCTTTGACAACCTTATCTGG 3412  
|||||  
QY 2743 GAGTTTTTCAGAACTGCTCTTCAAGTGTAGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 2802  
|||||  
Db 3413 GAGTTTTTCAGAACTGCTCTTCAAGTGTAGTTAGTTCAAGTGAACAGAGGCTGCTTCCAGTTT 3472  
|||||  
QY 2803 GGAGAAAAAGGAGGTTCCAGGAGTATGATTTAGCATTAAGTCAATTCGTAAGGAATCTTGG 2862  
|||||  
Db 3473 GGAGAAAAAGGAGGTTCCAGGAGTATGATTTTAGCATTAAGTCAATTCGTAAGGAATCTTGG 3532  
|||||  
QY 2863 ACTTGAGCACCTTAATGGATATATTTTACAGAGAACAGATCACTTTGGATGTATTTAGTTGA 2922  
|||||  
Db 3533 ACTTGAGCACCTTAATGGATATATTTTACAGAGAACAGATCACTTTGGATGTATTTAGTTGA 3592  
|||||  
QY 2923 GATGGGACACAGGAGCTGAAGAGATTTGGAATCAATGCTTATGACATAGGCACAAACT 2982  
|||||  
Db 3593 GATGGGACACAGGAGCTGAAGAGATTTGGAATCAATGCTTATGACATAGGCACAAACT 3652  
|||||  
QY 2983 AATTAAGGAGTGCAGAGACTTATCTCCGAGCAAAAGGCTTTACCCATATTTAACTTT 3042  
|||||

Db 3653 AATTAAGGAGTGCAGAGACTTATCTCCGAGCAACAAGGTCTTTAACCCATATTTAACTTT 3712  
QY 3043 GAACACCTCTGCTAGTGGAACTTCTTTATAGATCTGCTCTCTGATGATAAAGATTTCA 3102  
|||||  
Db 3713 GAACACCTCTGCTAGTGGAACTTCTTTATAGATCTGCTCTCTGATGATAAAGATTTCA 3772  
|||||  
QY 3103 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCACAGATGGAGTTCATGCAGG 3162  
|||||  
Db 3773 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCACAGATGGAGTTCATGCAGG 3832  
|||||  
QY 3163 TGGAACTTCTCAACAGATACAATATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTATG 3222  
|||||  
Db 3833 TGGAACTTCTCAACAGATACAATATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTATG 3892  
|||||  
QY 3223 GGAAGATACACTCACCGGAGAAAAAGTTTCTGAAGAAACCAACACCATGCAATGA 3282  
|||||  
Db 3893 GGAAGATACACTCACCGGAGAAAAAGTTTCTGAAGAAACCAACACCATGCAATGA 3952  
|||||  
QY 3283 AGCAATGCTTATTTTATGCTCTCTCTTTTGTGAATCAATTTATCCACAAGGCTTTGATGA 3342  
|||||  
Db 3953 AGCAATGCTTATTTTATGCTCTCTCTTTTGTGAATCAATTTATCCACAAGGCTTTGATGA 4012  
|||||  
QY 3343 AAGCATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGCTGAAAACCTCTTC 3402  
|||||  
Db 4013 AAGCATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGCTGAAAACCTCTTC 4072  
|||||  
QY 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGTCAGTTTCACAAAGA 3462  
|||||  
Db 4073 CAAAAGCAATCAATATGATATGGAATTTGGAGGAGGTACTGGGTGTCAGTTTCACAAAGA 4132  
|||||  
QY 3463 CAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 3522  
|||||  
Db 4133 CAGATCTTGTACATTTGCCACAGCAGCAGCTGCTCTTTTGGCGGTAACTTTGGGAAAGTC 4192  
|||||  
QY 3523 TTTCTCGAGTTCAGTGCATGCAATGAAATGGCAATTTCTCTCCAGGTCACTCAGTCAAC 3582  
|||||  
Db 4193 TTTCTCGAGTTCAGTGCATGCAATGAAATGGCAATTTCTCTCCAGGTCACTCAGTCAAC 4252  
|||||  
QY 3583 TGGTAGGCCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGAGAAC 3642  
|||||  
Db 4253 TGGTAGGCCCCAGTGTAAATGGCCTAGCATTAGCATTAGCTGAATATGTTATTTACAGAGAGAAC 4312  
|||||  
QY 3643 GGCTTATCTCAGTATTTTAAATTTACTTTACAGATTATGAGGCTGAAGGTATGGTCTGATGG 3702  
|||||  
Db 4313 GCTTATCTCAGTATTTTAAATTTACTTTACAGATTATGAGGCTGAAGGTATGGTCTGATGG 4372  
|||||  
QY 3703 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGG 3762  
|||||  
Db 4373 ATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGG 4432  
|||||  
QY 3763 CCTCTACGTTTACTCTCTTTGCTGAAAAAAA 3794  
|||||  
Db 4433 CCTCTACGTTTACTCTCTTTGCTGAAAAAAA 4464

## RESULT 8

AAF63952

ID AAF63952 standard; DNA: 5002 BP.

XX AAF63952;

AC AAF63952;

XX 05-APR-2001 (first entry)

DT Human tankyrase2 TANK2-LONG coding sequence SEQ ID NO: 132.

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder; ds.

XX Homo sapiens.

OS WO200100849-A1.

PN 04-JAN-2001.

XX

PD

XX 28-JUN-2000; 2000WO-US17827.  
 XX PF  
 XX XX  
 PR 29-JUN-1999; 99US-0141582.  
 XX XX  
 PA (ICOS-) ICOS CORP.  
 XX XX  
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 XX WPI; 2001-102896/11.  
 DR P-PSDB; AAB66294.  
 XX XX  
 PT New tankyrase2 polypeptides, useful for treating conditions mediated by  
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
 PT inflammatory and autoimmune disorders -  
 XX XX  
 PS Claim 5; Page 185-190; 242pp; English.  
 XX XX  
 CC The present invention provides the protein and coding sequence for the  
 CC human tankyrase2 protein. This is found in two different versions,  
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
 CC polyADP-ribosylation activity and is involved in the modification of  
 CC TRF1, which is a telomere-specific binding protein. The regulation of  
 CC telomere length, in which TRF1 has a role, is linked to ageing and  
 CC cancer. The sequences are useful in the treatment of cancers and  
 CC inflammatory disorders.  
 XX XX  
 SQ Sequence 5002 BP; 1456 A; 1065 C; 1196 G; 1285 T; 0 other;

Query Match 89.2%; Score 3387.2; DB 22; Length 5002;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 403 AGCTTTGGCGGAAAGACGTAGTGAATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC 462  
 DB 1083 AGGTTTGGCGGAAAGACGTAGTGAATTTGCTTCAGAAATGGTGCAAGTGTCCAAGC 1142  
 QY 463 ACCTGATGATGGGGGCTTATCTCTTCATATGATGCTCTTTTGGTCAATGCTGAAGT 522  
 DB 1143 ACCTGATGATGGGGGCTTATCTCTTCATATGATGCTCTTTTGGTCAATGCTGAAGT 1202  
 QY 523 AGTCAATCTCTTTTGGGACATGGTGACAGCCCAATGCTCGAGATAATTTGGAATTATAC 582  
 DB 1203 AGTCAATCTCTTTTGGGACATGGTGACAGCCCAATGCTCGAGATAATTTGGAATTATAC 1262  
 QY 583 TCCTCTCCATGAAGCTCAATTTAAAGAAAGATTGATGTTTGCATTTGCTGTTCACAGCA 642  
 DB 1263 TCCTCTCCATGAAGCTCAATTTAAAGAAAGATTGATGTTTGCATTTGCTGTTCACAGCA 1322  
 QY 643 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTTAGCAGATCC 702  
 DB 1323 TGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGACAGCATTTGGATTTAGCAGATCC 1382  
 QY 703 ATCTGCCAAAGCAGTCTTACTGTTGAATATAGAAAGATCAACTCTTAGAAAGTGCAG 762  
 DB 1383 ATCTGCCAAAGCAGTCTTACTGTTGAATATAGAAAGATCAACTCTTAGAAAGTGCAG 1442  
 QY 763 GAGTGGCAATGAAGAAATGATGGCTCTACTCACACCATTTAAATGTCACCTGCCACGC 822  
 DB 1443 GAGTGGCAATGAAGAAATGATGGCTCTACTCACACCATTTAAATGTCACCTGCCACGC 1502  
 QY 823 AAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATTAACAGAGTAAAGAT 882  
 DB 1503 AAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATTAACAGAGTAAAGAT 1562  
 QY 883 TGTACAGCTGTACTGCAACATGGAGCTGATGTCCTATTAAGATAAAGGTGATCTGGT 942  
 DB 1563 TGTACAGCTGTACTGCAACATGGAGCTGATGTCCTATTAAGATAAAGGTGATCTGGT 1622  
 QY 943 ACCATTACAAATGCCCTGTTCTTATGCTCAATATGAAGTAACTGAAGTAACTGTTGTCAGCA 1002  
 DB 1623 ACCATTACAAATGCCCTGTTCTTATGCTCAATATGAAGTAACTGAAGTAACTGTTGTCAGCA 1682

QY 1003 TGGTGCCTGTGTAATGCAATGGACTTTGGCAATTTCACTCTCTTCATGAGCAGCTTC 1062  
 DB 1683 TGGTGCCTGTGTAATGCAATGGACTTTGGCAATTTCACTCTCTTCATGAGCAGCTTC 1742  
 QY 1063 TAAGAACAGGGTTGAAGTATGTTCTTCTCTTTAAGTTATGGTGCAGACCCCAACACTGCT 1122  
 DB 1743 TAAGAACAGGGTTGAAGTATGTTCTTCTCTTTAAGTTATGGTGCAGACCCCAACACTGCT 1802  
 QY 1123 CAATTGTCAATAAAGTGTCTATAGACTTTGGCTCCACACACAGTAAAGAAAGATT 1182  
 DB 1803 CAATTGTCAATAAAGTGTCTATAGACTTTGGCTCCACACACAGTAAAGAAAGATT 1862  
 QY 1183 AGCATATGAATTTAAGAGCCACTCGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG 1242  
 DB 1863 AGCATATGAATTTAAGAGCCACTCGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCG 1922  
 QY 1243 AATCAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCCCTCAACACATGAAC 1302  
 DB 1923 AATCAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCCCTCAACACATGAAC 1382  
 QY 1303 AGCATTGCATTGTGCTGCTCATCTCCATATCCCAAGAAAGCAATATGTGAACCTGTT 1362  
 DB 1983 AGCATTGCATTGTGCTGCTCATCTCCATATCCCAAGAAAGCAATATGTGAACCTGTT 2042  
 QY 1363 GCTAAGAAAAGGAGCAACATCAATGAAAAGACTTAAAGAAATTTCTTGACTCTCTGCACT 1422  
 DB 2043 GCTAAGAAAAGGAGCAACATCAATGAAAAGACTTAAAGAAATTTCTTGACTCTCTGCACT 2102  
 QY 1423 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGTGTTGAAACATGAAGCAAGGT 1482  
 DB 2103 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGTGTTGAAACATGAAGCAAGGT 2162  
 QY 1483 TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTCACT 1542  
 DB 2163 TAATGCTCTGGATAAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTCACT 2222  
 QY 1543 ACAAACCTGCCGCTACTCTGAGCTATGGGTGTGATCCTAACATTTATATCCCTTCAGGG 1602  
 DB 2223 ACAAACCTGCCGCTACTCTGAGCTATGGGTGTGATCCTAACATTTATATCCCTTCAGGG 2282  
 QY 1603 CTTTACTGCTTTACAGATGGAAATGAAATGTACACCAACTCTCCAAAGAGGTATCTC 1662  
 DB 2283 CTTTACTGCTTTACAGATGGAAATGAAATGTACACCAACTCTCCAAAGAGGTATCTC 2342  
 QY 1663 ATTAGTAAATTCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAAAGCTGGAGATGCGA 1722  
 DB 2343 ATTAGTAAATTCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAAAGCTGGAGATGCGA 2402  
 QY 1723 AACTGTAAAAAACTGTGTACTGTTACAGAGTCAACTGCAGAGACATTTGAAGGCGCTCA 1782  
 DB 2403 AACTGTAAAAAACTGTGTACTGTTACAGAGTCAACTGCAGAGACATTTGAAGGCGCTCA 2462  
 QY 1783 GTCTACACCACTTCTATTTGTCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT 1842  
 DB 2463 GTCTACACCACTTCTATTTGTCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCT 2522  
 QY 1843 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACTTTTGCACATGC 1902  
 DB 2523 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACTTTTGCACATGC 2582  
 QY 1903 ATGTTCTTATGGACATTTGAAGTGTGAGACTTCTTGTAAACATGAGCAGTAGTTAA 1962  
 DB 2583 ATGTTCTTATGGACATTTGAAGTGTGAGACTTCTTGTAAACATGAGCAGTAGTTAA 2642  
 QY 1963 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGA 2022  
 DB 2643 TGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGA 2702  
 QY 2023 AATTGTCAAACTCTGCTCCAGCATGTTGTCAGACCCCTACCAAAAAAACACAGGGATGAAA 2082  
 DB 2703 AATTGTCAAACTCTGCTCCAGCATGTTGTCAGACCCCTACCAAAAAAACACAGGGATGAAA 2762  
 QY 2083 TACTCCCTTTGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAGA 2142

|||||  
Db 2763 TACTCCTTTGGAGTCTGTTAAAGATGGAGATACAGATATCAAGATCTGCTTAGGGGAGA 2822  
|||||  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGTTGTTATCCAGAGTGAAGAAGTTGTCCTC 2202  
|||||  
Db 2823 TGCAGCTTTGCTAGATGCTGCCAAGAAGGTTGTTATCCAGAGTGAAGAAGTTGTCCTC 2882  
|||||  
QY 2203 TCGTATAATGTAAATTTGCCGCGATACCAAGCGAGACATTCACACACCTTTACATTTAGC 2262  
|||||  
Db 2883 TCCTGATATGTAAATTTGCCGCGATACCAAGCGAGACATTCACACACCTTTACATTTAGC 2942  
|||||  
QY 2263 AGCTGGTTATATAATTTAGAAGTTGCCAGATTTGTTTACAACACGAGGCTGATGTCAA 2322  
|||||  
Db 2943 AGCTGGTTATATAATTTAGAAGTTGCCAGATTTGTTTACAACACGAGGCTGATGTCAA 3002  
|||||  
QY 2323 TGCCCAAGACAAGGAGGACTTATTCCTTTACATATATGCAGCATCTTACGGGCATGTAGA 2382  
|||||  
Db 3003 TGCCCAAGACAAGGAGGACTTATTCCTTTACATATATGCAGCATCTTACGGGCATGTAGA 3062  
|||||  
QY 2383 TGTAGCAGCTCTACTAAATAAAGTATAATGCTGTGTCATGCGCAGGACAAATGGGCTTT 2442  
|||||  
Db 3063 TGTAGCAGCTCTACTAAATAAAGTATAATGCTGTGTCATGCGCAGGACAAATGGGCTTT 3122  
|||||  
QY 2443 CACACCTTTGCGAAGACAGGCCAAAGGAGCAACACAGCTTTGTGCTTGTGCTAGC 2502  
|||||  
Db 3123 CACACCTTTGCGAAGACAGGCCAAAGGAGCAACACAGCTTTGTGCTTGTGCTAGC 3182  
|||||  
QY 2503 CCATGGAGCTGACCCGACTCTTTAAATAATCAGGAAGCAACACACTTTAGATTAGTTTC 2562  
|||||  
Db 3183 CCATGGAGCTGACCCGACTCTTTAAATAATCAGGAAGCAACACACTTTAGATTAGTTTC 3242  
|||||  
QY 2563 AGCGATGATGTACAGCGCTCTTCTGACAGCAGCCATGCCCCATCTGCTGCCCCCTTTG 2622  
|||||  
Db 3243 AGCAGATGATGTACAGCGCTCTTCTGACAGCAGCCATGCCCCATCTGCTGCCCCCTTTG 3302  
|||||  
QY 2623 TTACAAGCTCAAGTGCTCAATGGTGTGAGAAGCCGAGAGCCACTGCAGATGCTCTCTC 2682  
|||||  
Db 3303 TTACAAGCTCAAGTGCTCAATGGTGTGAGAAGCCGAGAGCCACTGCAGATGCTCTCTC 3362  
|||||  
QY 2683 TTTACAGTCCATCTAGCCCATCAAGCTTTCTGACAGCAGCCTTTCAGCCAGCAGCTTATCTGG 2742  
|||||  
Db 3363 TTTACAGTCCATCTAGCCCATCAAGCTTTCTGACGCGAGCAGCTTTGACAACTTATCTGG 3422  
|||||  
QY 2743 GAGTTTTTTCAGAACTGCTCTCAGTAGTTAGTTCAAGTGAAGCAGAGGCTGCTTCCAGTTT 2802  
|||||  
Db 3423 GAGTTTTTTCAGAACTGCTCTCAGTAGTTAGTTCAAGTGAAGCAGAGGCTGCTTCCAGTTT 3482  
|||||  
QY 2803 GGAGAAAAAGGAGTTCCAGGAGTAGATTTAGCATAACTCAATTCGTAAGGAATCTTTGG 2862  
|||||  
Db 3483 GGAGAAAAAGGAGTTCCAGGAGTAGATTTAGCATAACTCAATTCGTAAGGAATCTTTGG 3542  
|||||  
QY 2863 ACTTGAGCACCTAATGGATATATTTGAGAGACAAGATCACTTTGGATGTATTAGTTGA 2922  
|||||  
Db 3543 ACTTGAGCACCTAATGGATATATTTGAGAGACAAGATCACTTTGGATGTATTAGTTGA 3602  
|||||  
QY 2923 GATGGGCAACAAGGAGCTGAAGGAGATTCGAATCAATGCTTATGACATAGGCACAAACT 2982  
|||||  
Db 3603 GATGGGCAACAAGGAGCTGAAGGAGATTCGAATCAATGCTTATGACATAGGCACAAACT 3662  
|||||  
QY 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGACAACAAGGCTTTAACCCATATTTAACTTT 3042  
|||||  
Db 3663 AATTAAAGGAGTCGAGAGACTTATCTCCGACAACAAGGCTTTAACCCATATTTAACTTT 3722  
|||||  
QY 3043 GAACACCTCTGTTAGTGGAACTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTCA 3102  
|||||  
Db 3723 GAACACCTCTGTTAGTGGAACTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTCA 3782  
|||||  
QY 3103 GTCGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCACAGATGGAGGTCATGCAGG 3162  
|||||  
Db 3783 GTCGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCACAGATGGAGGTCATGCAGG 3842  
|||||  
QY 3163 TGGAACTTTCAACAGATACAATATTTCTCAAGATTTCAGAGGTTTCTGAACAAGAACTATG 3222  
|||||

Db 3843 TGGAACTTTCAACAGATACAATATTTCTCAAGATTTCAGAAGGTTTGTAAACAAGAACTATG 3902  
QY 3223 GGAAGATACACTCACCGAGAAAAGTTCCTGAGAAGAAACACACACCATGCCAATGA 3282  
|||||  
Db 3903 GGAAGATACACTCACCGAGAAAAGTTCCTGAGAAGAAACACACACCATGCCAATGA 3962  
|||||  
QY 3283 ACGAATGCTATTTTCATGGGCTCTCCTTTTGTGAATGCAATTTATCCACAAGGCTTTTGATGA 3342  
|||||  
Db 3963 ACGAATGCTATTTTCATGGGCTCTCCTTTTGTGAATGCAATTTATCCACAAGGCTTTTGATGA 4022  
|||||  
QY 3343 AAGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTCTGCTGAAACTCTTC 3402  
|||||  
Db 4023 AAGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTCTGCTGAAACTCTTC 4082  
|||||  
QY 3403 CAAAGCAATCAATATATATGAAATTTGGAGGAGTACTGGGTGTCAGTTCCACAAGA 3462  
|||||  
Db 4083 CAAAGCAATCAATATATGAAATTTGGAGGAGTACTGGGTGTCAGTTCCACAAGA 4142  
|||||  
QY 3463 CAGATCTTGTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACTTTGGGAAAGTC 3522  
|||||  
Db 4143 CAGATCTTGTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACTTTGGGAAAGTC 4202  
|||||  
QY 3523 TTTCTCTGAGTTTCAGTGCATGAAATGGCACATCTCTCTCCAGGTCATCCTCAGTCAC 3582  
|||||  
Db 4203 TTTCTCTGAGTTTCAGTGCATGAAATGGCACATCTCTCTCCAGGTCATCCTCAGTCAC 4262  
|||||  
QY 3583 TGGTAGCCCAAGTAAATGSCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642  
|||||  
Db 4263 TGGTAGCCCAAGTAAATGSCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 4322  
|||||  
QY 3643 GGCATTCTCTGAGTATTTAATTACTTACCAGATTATGAGGCTGAAGGTATGGTCTGATGG 3702  
|||||  
Db 4323 GGCATTCTCTGAGTATTTAATTACTTACCAGATTATGAGGCTGAAGGTATGGTCTGATGG 4382  
|||||  
QY 3703 ATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGG 3762  
|||||  
Db 4383 ATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGG 4442  
|||||  
QY 3763 CCTCTAGTTTTACTCTCTTCTGCTGAAAAAAA 3794  
|||||  
Db 4443 CCTCTAGTTTTACTCTCTTCTGCTGAAAAAAA 4474  
|||||  
RESULT 9  
AAD02578  
ID AAD02578 standard; DNA; 4512 BP.  
XX  
AC AAD02578;  
XX  
DT 02-MAY-2001 (first entry)  
XX  
DE Human tankyrase homolog protein (THP) DNA.  
XX  
KW Human; tankyrase homolog protein; THP; gene therapy; cancer;  
KW tumour; basal cell carcinoma; therapy; genetic mapping;  
KW cytoskeletal; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 23..3523  
FT /\*tag= a  
FT /product= "Human tankyrase homolog protein (THP)"  
FT /note= "This region is specifically claimed in  
claim 4 as SEQ ID NO:4"  
PN WO200104326-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 03-JUL-2000; 2000WO-EF06609.  
XX  
PR 09-JUL-1999; 99US-0350982.

XX (PHAA ) PHARMACIA & UPJOHN SPA.  
PA Berthelsen J, Toma S, Isacchi A;  
PI  
XX  
XX  
DR WPI: 2001-168422/17.  
DR P-PSDB; AAY72589.  
XX  
XX  
PT New tankyrase homolog protein (THP) polynucleotide and polypeptide  
PT useful in gene therapy, diagnosis and treatment or prevention of  
PT unregulated cell growth, such as cancer or tumor cell growth -  
XX  
XX  
PS Claim 1; Page 41-42; 60pp; English.  
XX  
CC The present sequence is a DNA coding for human tankyrase homolog  
CC protein (THP). The THP polypeptides and polynucleotides of the  
CC invention are useful in gene therapy and for treating or preventing  
CC unregulated cell growth such as cancer or tumour (e.g. basal cell  
CC carcinoma). The nucleic acid molecules of the invention and their  
CC fragments are useful for restriction fragment length polymorphism  
CC (RFLP) associated with certain disorders, as well as for genetic  
CC mapping. Antisense oligonucleotides, or fragments of nucleic acid  
CC encoding THP are useful as diagnostic tools for probing the  
CC expression of Thp gene in various tissues. THP can be used as  
CC antigens for raising antibodies against them and in assays for  
CC identifying compounds that modulate their activity. They are used in  
CC the manufacture of a medicament directed towards cancers or tumours.  
CC THP are also useful for screening compounds in a variety of drugs  
CC screening techniques and as a research tool for identification,  
CC characterisation and purification of interacting, regulatory proteins.  
XX  
SQ Sequence 4512 BP; 1356 A; 878 C; 1002 G; 1268 T; 8 other;

Query Match 89.1%; Score 3382.8; DB 22; Length 4512;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3381; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 403 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTCCAAAGTCCCAAGC 462  
DB |||||  
DB 220 AGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAATGGTCCAAAGTCCCAAGC 279  
QY 463 ACCTGATGATGGGGCTTATTCCTCTTCATAATGATGCTCTTTGGTCATGCTGAAGT 522  
DB |||||  
DB 280 ACCTGATGATGGGGCTTATTCCTCTTCATAATGATGCTCTTTGGTCATGCTGAAGT 339  
QY 523 AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAAATGGAAATTATAC 582  
DB |||||  
DB 340 AGTCAATCTCTTTTGGACATGGTGCAGACCCCAATGCTCGAGATAAATGGAAATTATAC 399  
QY 583 TCCTCTCCATGAAAGCTGCAATTAAGGAAAGATTGATGTTTGCAATGCTGTGTACAGCA 642  
DB |||||  
DB 400 TCCTCTCCATGAAAGCTGCAATTAAGGAAAGATTGATGTTTGCAATGCTGTGTACAGCA 459  
QY 643 TGGAGCTGAGCCCAACCATCGCAATACAGATGGAAGACAGCAATGGATTTACAGATCC 702  
DB |||||  
DB 460 TGGAGCTGAGCCCAACCATCGCAATACAGATGGAAGACAGCAATGGATTTACAGATCC 519  
QY 703 ATCTGCAAGACGAGTCTTACTGCTGAATATAAGAAAGATGAAGTCTTAGAAAGTGCAG 762  
DB |||||  
DB 520 ATCTGCAAGACGAGTCTTACTGCTGAATATAAGAAAGATGAAGTCTTAGAAAGTGCAG 579  
QY 763 GAGTGGCAATGAAGAAAAATGATGCTCTTACTCACACCAATTAATGTCAACTGCCACGC 822  
DB |||||  
DB 580 GAGTGGCAATGAAGAAAAATGATGCTCTTACTCACACCAATTAATGTCAACTGCCACGC 639  
QY 823 AAGTGTGCGAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGATAAGAT 882  
DB |||||  
DB 640 AAGTGTGCGAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGATAAGAT 699  
QY 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTTAAAGATAAAGGTGATCTGGT 942  
DB |||||  
DB 700 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTTAAAGATAAAGGTGATCTGGT 759

QY 943 ACCATTACACAATGCCCTGTTCTTTATGGTCATTATGAAGTAACCTGAACCTTTTGGTCAAGCA 1002  
DB |||||  
DB 760 ACCATTACACAATGCCCTGTTCTTTATGGTCATTATGAAGTAACCTTTTGGTCAAGCA 819  
QY 1003 TGGTGCCCTGTAAATGCAATGGACTTGTGGCAATTCACCTCTTTCATGAGGAGCTTC 1062  
DB |||||  
DB 820 TGGTGCCCTGTAAATGCAATGGACTTGTGGCAATTCACCTCTTTCATGAGGAGCTTC 879  
QY 1063 TAGAACACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTCAGACCCCAACACTGCT 1122  
DB |||||  
DB 880 TAGAACACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGGTCAGACCCCAACACTGCT 939  
QY 1123 CAATTGTGCACATAAAAGTCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1182  
DB |||||  
DB 940 CAATTGTGCACATAAAAGTCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 999  
QY 1183 AGCATATGAATTTAAGGCCACCTCGTTGCTGCAAGCTGCACGAGAACTGATGTTACTCG 1242  
DB |||||  
DB 1000 AGCATATGAATTTAAGGCCACCTCGTTGCTGCAAGCTGCACGAGAACTGATGTTACTCG 1059  
QY 1243 AATCAAAAAACATCTCTCTCTGGAATGGTGAATTTTCAAGCATCCTCAACACATGAAC 1302  
DB |||||  
DB 1060 AATCAAAAAACATCTCTCTCTGGAATGGTGAATTTTCAAGCATCCTCAACACATGAAC 1119  
QY 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAAATATCTGAACCTGTT 1362  
DB |||||  
DB 1120 AGCANTGCATTGTGCTGCTCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAACCTGTT 1179  
QY 1363 GCTAAGAAAGAGGACAAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGACGT 1422  
DB |||||  
DB 1180 GCTAAGAAAGAGGACAAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGACGT 1239  
QY 1423 GGCATCTGAGAAAGCTCATATATGATGTTGTTGAAGTAGTGTGAAACATCAAGCAAAAGT 1482  
DB |||||  
DB 1240 GGCATCTGAGAAAGCTCATATATGATGTTGTTGAAGTAGTGTGAAACATCAAGCAAAAGT 1299  
QY 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTTGCTCATCT 1542  
DB |||||  
DB 1300 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTTGCTCATCT 1359  
QY 1543 ACAAACTGCCGCTTACTCTGAGCTATGGGTGTGATCCTTAACATATATATCCCTTTCAGG 1602  
DB |||||  
DB 1360 ACAAACTGCCGCTTACTCTGAGCTATGGGTGTGATCCTTAACATATATATCCCTTTCAGG 1419  
QY 1603 CTTTACTGCTTTACAGATGGCAATGAAATGTACAGCAACTCCTCCAAGAGGATATCTC 1662  
DB |||||  
DB 1420 CTTTACTGCTTTACAGATGGCAATGAAATGTACAGCAACTCCTCCAAGAGGATATCTC 1479  
QY 1663 ATTAGGTAATTCAGAGCGACAGACAAATTCCTGGAAGCTGCAAAAGGCTGGAGATGTCGA 1722  
DB |||||  
DB 1480 ATTAGGTAATTCAGAGCGACAGACAAATTCCTGGAAGCTGCAAAAGGCTGGAGATGTCGA 1539  
QY 1723 AACTGTAAAAAACTGTGACTGTTGAGAGTGTCACTGAGAGACATTTGAAGGGCTGCA 1782  
DB |||||  
DB 1540 AACTGTAAAAAACTGTGACTGTTGAGAGTGTCACTGAGAGACATTTGAAGGGCTGCA 1599  
QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTGCTGCTGGAATATCTGCT 1842  
DB |||||  
DB 1600 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTGCTGCTGGAATATCTGCT 1659  
QY 1843 ACAGCATGGAGCTGATGTGCATCTTAAAGATAAAGGAGGCTTTGTACCTTTTGCACAATGC 1902  
DB |||||  
DB 1660 ACAGCATGGAGCTGATGTGCATCTTAAAGATAAAGRRGSCCTTTGTACCTTTTGCACAATGC 1719  
QY 1903 ATGTTCTTATGGACATATGAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAA 1962  
DB |||||  
DB 1720 ATGTTCTTATGGACATATGAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAA 1779  
QY 1963 TGTAGCTGATTTATGGAATTTTACACTTTTACATGAAGCAGCAAAAGCAAAATATGA 2022  
DB |||||  
DB 1780 TGTAGCTGATTTATGGAATTTTACACTTTTACATGAAGCAGCAAAAGCAAAATATGA 1839  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTTACCACAAAAAACAGGATGGAAA 2082



|||||  
Db 1840 AATTGGCAACTCTGCTCCAGCATGTCGAGACCTTACAAAAACACAGGATGAAA 1899  
Qy 2083 TACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCACAGATCTGCTTAGGGGAGA 2142  
Db 1900 TACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCACAGATCTGCTTAGGGGAGA 1959  
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTGAAGAGTTGCTTC 2202  
Db 1960 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTACGAGAGTGAAGAGTTGCTTC 2019  
Qy 2203 TCCTGATAATGTAATTTGCCGCGATACCCAGGCGAGACATTCACACCTTTTACATTTAGC 2262  
Db 2020 TCCTGATAATGTAATTTGCCGCGATACCCAGGCGAGACATTCACACCTTTTACATTTAGC 2079  
Qy 2263 AGCTGGTTATTAATTTAGAGTTGACAGATATTTGTTTACACACAGGAGCTGATGTGA 2322  
Db 2080 AGCTGGTTATTAATTTAGAGTTGACAGATATTTGTTTACACACAGGAGCTGATGTGA 2139  
Qy 2323 TGCCCAAGACAAAGAGGACTTATTCCTTTACATAATGACAGCATCTTACGGGATGTAGA 2382  
Db 2140 TGCCCAAGACAAAGAGGACTTATTCCTTTACATAATGACAGCATCTTACGGGATGTAGA 2199  
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGTCATGTGTCAATGCCACGACAAATGGGCTTT 2442  
Db 2200 TGTAGCAGCTCTACTAATAAAGTATAATGTCATGTGTCAATGCCACGACAAATGGGCTTT 2259  
Qy 2443 CACACCTTTGACAGACAGCCCAAGAGGAGACACAGCTTTGCTTTGTCGTAGC 2502  
Db 2260 CACACCTTTGACAGACAGCCCAAGAGGAGACACAGCTTTGCTTTGTCGTAGC 2319  
Qy 2503 CCATGGAGCTGACCGGACTCTTAAATATCAGGAAGGACAAACACTTTTAGATTTAGTTTC 2562  
Db 2320 CCATGGAGCTGACCGGACTCTTAAATATCAGGAAGGACAAACACTTTTAGATTTAGTTTC 2379  
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622  
Db 2380 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2439  
Qy 2623 TTACAGCCTCAAGTCTCAATGGTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTTC 2682  
Db 2440 TTACAGCCTCAAGTCTCAATGGTGTGAGAGAGCCAGGAGCCACTGCAGATGCTCTTC 2499  
Qy 2683 TTCAGGTCATCTAGCCCATCAAGCCTTTCTGACAGCAGCTCTGACAACTTTATCTGG 2742  
Db 2500 TTCAGGTCATCTAGCCCATCAAGCCTTTCTGACAGCAGCTCTGACAACTTTATCTGG 2559  
Qy 2743 GAGTTTTTCAGAACTGTCTTCCAGTAGTTTGTCAAGTGAAGAGGAGGCTTCCAGTTT 2802  
Db 2560 GAGTTTTTCAGAACTGTCTTCCAGTAGTTTGTCAAGTGAAGAGGAGGCTTCCAGTTT 2619  
Qy 2803 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATACTCAATTCGTAAGAACTTTGG 2862  
Db 2620 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATACTCAATTCGTAAGAACTTTGG 2679  
Qy 2863 ACTTGACACCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGA 2922  
Db 2680 ACTTGACACCTTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGATTAGTTGA 2739  
Qy 2923 GATGGGACAAAGAGGCTGAAGAGGATTTGGAATCAATGCTTATGACATAGGACACAACT 2982  
Db 2740 GATGGGACAAAGAGGCTGAAGAGGATTTGGAATCAATGCTTATGACATAGGACACAACT 2799  
Qy 2983 AATTAAAGAGTCGAGAGACTTATCTCCGACACAAAGGCTTAACCCATATTTAACTTT 3042  
Db 2800 AATTAAAGAGTCGAGAGACTTATCTCCGACACAAAGGCTTAACCCATATTTAACTTT 2859  
Qy 3043 GAACACCTCTGCTAGTGAACAAATTTTATAGATCTGCTCTGATGATAAAGAGTTTCA 3102  
Db 2860 GAACACCTCTGCTAGTGAACAAATTTTATAGATCTGCTCTGATGATAAAGAGTTTCA 2919  
Qy 3103 GTCTGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGAGATGGAGGTCATGTCAGG 3162  
|||||

Db 2920 GTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCAGAGATGGAGGTCATCCAGG 2979  
Qy 3163 TGGAAATCTTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTATG 3222  
Db 2980 TGGAAATCTTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTATG 3039  
Qy 3223 GGAAGATACACTCACCGGAGAAAAGTTTCTGAAAGAAAACCAACCATGCCAATGA 3282  
Db 3040 GGAAGATACACTCACCGGAGAAAAGTTTCTGAAAGAAAACCAACCATGCCAATGA 3099  
Qy 3283 ACGAATGCTATTTCATGGGTCCTCTTTGTGAATGCAATATTCACAAAGGCTTTGATGA 3342  
Db 3100 ACGAATGCTATTTCATGGGTCCTCTTTGTGAATGCAATATTCACAAAGGCTTTGATGA 3159  
Qy 3343 AAGGCATGCTACATAGGTGGTATGTTTGGAGCTGCAATTTATTTGCTGAAAACCTCTTC 3402  
Db 3160 AAGGCATGCTACATAGGTGGTATGTTTGGAGCTGCAATTTATTTGCTGAAAACCTCTTC 3219  
Qy 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTGTCACAAAGA 3462  
Db 3220 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGTACTGGGTGTCACAAAGA 3279  
Qy 3463 CAGATCTTGTAGATTTGCCACAGGAGCTGCTCTTTTCCGGGTAACTTTGGGAAGTC 3522  
Db 3280 CAGATCTTGTAGATTTGCCACAGGAGCTGCTCTTTTCCGGGTAACTTTGGGAAGTC 3339  
Qy 3523 TTTCTGTCAGTTTCAGTGAATGGAATGGCAATTTCCCTCCAGGTCATCAGTCTAC 3582  
Db 3340 TTTCTGTCAGTTTCAGTGAATGGAATGGCAATTTCCCTCCAGGTCATCAGTCTAC 3399  
Qy 3583 TGTAGGCCCAAGTGAATGGCTAGCATTTAGTGAATATGTTATTTACAGAGGAGAA 3642  
Db 3400 TGTAGGCCCAAGTGAATGGCTAGCATTTAGTGAATATGTTATTTACAGAGGAGAA 3459  
Qy 3643 GGCTTATCTGAGATTTAATTTACTTTACAGATTTAGGCCCTGAAGGCTATGGTTCATGG 3702  
Db 3460 GGCTTATCTGAGATTTAATTTACTTTACAGATTTAGGCCCTGAAGGCTATGGTTCATGG 3519  
Qy 3703 ATAAATAGTTATTTTAAAGAACTAAATTCACACTGAACCTCAAAAGCAGCAGTGG 3762  
Db 3520 ATAAATAGTTATTTTAAAGAACTAAATTCACACTGAACCTCAAAAGCAGCAGTGG 3579  
Qy 3763 CCTCTAGTTTTACTCTCTTTGCTGAAAAAAA 3794  
Db 3580 CCTCTAGTTTTACTCTCTTTGCTGAAAAAAA 3611  
RESULT 10  
AA25366  
ID AAX25366 standard; cdna; 3400 BP.  
XX  
AC AAX25366;  
XX  
DT 19-JUL-1999 (first entry)  
XX  
DE Human Grb7 effector 2.2412 cdna.  
XX  
KW Grb7 effector; 2.2412 protein; human; signal transduction;  
tumour marker; breast cancer; prostate cancer; prognosis;  
diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT CDS 1..3225  
FT /Lag= a  
XX  
XX WO9915647-A1.  
XX  
XX PD 01-APR-1999.  
XX  
XX 23-SEP-1998; 98WO-AU00795.  
XX

PR 23-SEP-1997; 97AU-0009388.  
 XX (GARV-) GARVAN INST MEDICAL RES.  
 XX Daily RJ, Sutherland RL;  
 XX WPI; 1999-254707/21.  
 DR P-PSDB; AAY05734.  
 XX  
 PT New candidate effector for the Grb7 family of signaling proteins,  
 PT and specific antibody, useful for detection and treatment of cancer  
 XX  
 PS Claim 4; Fig 1; 24pp; English.  
 XX  
 CC This is the nucleotide sequence of an isolated polynucleotide  
 CC molecule encoding a candidate effector protein, termed 2.2412 (see  
 CC AAY05734), for the Grb7 family of signalling proteins. A partial  
 CC clone was isolated from a human liver cDNA library using a yeast  
 CC two-hybrid system assay with Grb14 as bait. This clone was used  
 CC to screen a human placental cDNA library that provided the  
 CC c-terminal end of the 2.2412 sequence. A 5' sequence for the  
 CC clone, including the initiation codon, was not identified.  
 CC Analysis of the sequence revealed significant homology to a large  
 CC number of proteins containing ankyrin-like repeats. The gene  
 CC was localised to between chromosome 10q23.2 and proximal 10q23.32.  
 CC Deletions in the 10q22-25 region have been detected in human breast,  
 CC prostate, renal, small cell lung and endometrial carcinomas,  
 CC glioblastoma multiforme, melanoma and meningiomas. Detection of  
 CC the protein encoded by the 2.2412 cDNA in a sample should provide a  
 CC useful tumour marker and/or prognostic indicator for certain human  
 CC cancers, in particular breast cancer and prostate cancer.  
 CC Antagonism of the interaction between Grb7 family members and the  
 CC encoded protein should provide a novel treatment strategy for human  
 CC diseases exhibiting aberrant receptor tyrosine kinase signalling,  
 CC such as cancer. Oligonucleotide probes can be used in methods of  
 CC detecting the presence of 2.2412 mRNA in a sample.  
 XX  
 SQ Sequence 3400 BP; 1059 A; 565 G; 760 G; 916 T; 0 other;

Query Match 86.8%; Score 3297; DB 20; Length 3400;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 482 ATTCCCTTTCAATGAATGCTGCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGGCA 541  
 DB 1 ATTCCCTTTCAATGAATGCTGCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGGCA 60  
 QY 542 CATGTGTCAGACCCCAATGCTCGAGATAATTGGAAATTAATACCTCTCCATGAAGCTGCA 601  
 DB 61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAAATTAATACCTCTCCATGAAGCTGCA 120  
 QY 602 ATTAAGGAAAGATGATGTTTGGCAATGCTGCTGTTACAGCATGGAGCTGAGCCCAACCATC 661  
 DB 121 ATTAAGGAAAGATGATGTTTGGCAATGCTGCTGTTACAGCATGGAGCTGAGCCCAACCATC 180  
 QY 662 CGAAATACAGATGGAAGACACATGATTTAGCAGATCCATCTGCCAAAGCAGTGTCT 721  
 DB 181 CGAAATACAGATGGAAGACACATGATTTAGCAGATCCATCTGCCAAAGCAGTGTCT 240  
 QY 722 ACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCAGAGTGGCAATGAAGAAAA 781  
 DB 241 ACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCAGAGTGGCAATGAAGAAAA 300  
 QY 782 ATGATGGCTCTACTCACACCATTAATGTCACTGCCACGCAAGTGATGGCAGAAAGTCA 841  
 DB 301 ATGATGGCTCTACTCACACCATTAATGTCACTGCCACGCAAGTGATGGCAGAAAGTCA 360  
 QY 842 ACTCCATTACATTTTGGCAGCAGGATATAACAGACTTAAGATTTACACCTGTTACTGCAA 901  
 DB 361 ACTCCATTACATTTTGGCAGCAGGATATAACAGACTTAAGATTTACACCTGTTACTGCAA 420  
 QY 902 CATGGAGCTGATGTCATGCTTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCCTGT 961  
 DB 902 CATGGAGCTGATGTCATGCTTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCCTGT 961

DB 421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCCTGT 480  
 QY 962 TCTTATGCTCAITATGAAGTAACCTGAACCTTTTGGTCAAGCATGGTGTGCTGTAATGCA 1021  
 DB 481 TCTTATGCTCAITATGAAGTAACCTGAACCTTTTGGTCAAGCATGGTGTGCTGTAATGCA 540  
 QY 1022 ATGCACTTGTGGCAATTCACCTCTTTCATGAGCAGCTTCTTAAGAACAGGGTTGAAGTA 1081  
 DB 541 ATGCACTTGTGGCAATTCACCTCTTTCATGAGCAGCTTCTTAAGAACAGGGTTGAAGTA 600  
 QY 1082 TGTTCCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGCTCAATTTGCAATAAAGT 1141  
 DB 601 TGTTCCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGCTCAATTTGCAATAAAGT 660  
 QY 1142 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 1201  
 DB 661 GCTATAGACTTGGCTCCACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 720  
 QY 1202 CACTCGTGTGCTGCAAGTGCACGAGAGAGCTGATCTTACTCGAATCAAAAACATCTCTCT 1261  
 DB 721 CACTCGTGTGCTGCAAGTGCACGAGAGAGCTGATCTTACTCGAATCAAAAACATCTCTCT 780  
 QY 1262 CTGAAATGCTGCAATTTCAAGCATCTCAACACATGAAGAGAGCTTGTGCTGCT 1321  
 DB 781 CTGAAATGCTGCAATTTCAAGCATCTCAACACATGAAGAGAGCTTGTGCTGCT 840  
 QY 1322 GCATCTCCATATCCCAAGAAAGCAATATGCAACTGTTGCTAAGAAAGAGCAAAAC 1381  
 DB 841 GCATCTCCATATCCCAAGAAAGCAATATGCAACTGTTGCTAAGAAAGAGCAAAAC 900  
 QY 1382 ATCAATGAAAGAGCTAAAGAAATTTGACTCTCTGACCTGGCATCTGAGAAGCTCAT 1441  
 DB 901 ATCAATGAAAGAGCTAAAGAAATTTGACTCTCTGACCTGGCATCTGAGAAGCTCAT 960  
 QY 1442 AATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGTTAATGCTGCGATAATCTT 1501  
 DB 961 AATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGGTTAATGCTGCGATAATCTT 1020  
 QY 1502 GGTGAGACTTCTCTACAGAGCTGCAATATTTGGTCACTATCAAAACCTGCCCTACTC 1561  
 DB 1021 GGTGAGACTTCTCTACAGAGCTGCAATATTTGGTCACTATCAAAACCTGCCCTACTC 1080  
 QY 1562 CTGAGCTATGGTGTGATCCTAACATTTATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621  
 DB 1081 CTGAGCTATGGTGTGATCCTAACATTTATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140  
 QY 1622 GGAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATTTAGTATTCAGAGGCA 1681  
 DB 1141 GGAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTCATTTAGTATTCAGAGGCA 1200  
 QY 1682 GACAGCAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCGAAACTGTAAAAAAGCTGTGT 1741  
 DB 1201 GACAGCAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCGAAACTGTAAAAAAGCTGTGT 1260  
 QY 1742 ACTGTTTACAGAGTGTCACTGACAGACATGAAGGGCGTCAGTCTACACACTTTCATTTT 1801  
 DB 1261 ACTGTTTACAGAGTGTCACTGACAGACATGAAGGGCGTCAGTCTACACACTTTCATTTT 1320  
 QY 1802 GCAGCTGGGTATACAGAGTGTCCGTTGGGAATATCTGCTACAGCATGAGCTGATGTG 1861  
 DB 1321 GCAGCTGGGTATACAGAGTGTCCGTTGGGAATATCTGCTACAGCATGAGCTGATGTG 1380  
 QY 1862 CATGCTAAAGATAAAGGAGGCTTTGACCTTTGCACAAATGCAATGCTTCTTATGGACATTTAT 1921  
 DB 1381 CATGCTAAAGATAAAGGAGGCTTTGACCTTTGCACAAATGCAATGCTTCTTACGACATTTAT 1440  
 QY 1922 GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTAAATGCTAGCTGATTTATGGAAA 1981  
 DB 1441 GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTAAATGCTAGCTGATTTATGGAAA 1500  
 QY 1982 TTTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTCTGCTC 2041  
 DB 1501 TTTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTCTGCTC 1560

QY 2042 CAGCATGGTCAGACCCCTACCAAAAAAACAAGGATCGAAATACTCCTTTGGATCTTGTT 2101  
Db 1561 CAGGATGGTCAGACCCCTACAAAAAACAAGGATCGAAATACTCCTTTGGATCTTGTT 1620  
QY 2102 AAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAGATGAGCTTTGCTAGATGCT 2161  
Db 1621 AAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAGATGAGCTTTGCTAGATGCT 1680  
QY 2162 GCCAAGAGGGTGTTCAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATGTC 2221  
Db 1681 GCCAAGAGGGTGTTCAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATGTC 1740  
QY 2222 CGCGATCCCAAGGACAGATTCACACCTTTACATTTAGCAGCTGCTTATAATTTA 2281  
Db 1741 CGCGATCCCAAGGACAGATTCACACCTTTACATTTAGCAGCTGCTTATAATTTA 1800  
QY 2282 GAAAGTGCAGAGTATTTGTTACACAGGAGCTGATGTGAATGCCAAGCAAAAGGAGGA 2341  
Db 1801 GAAAGTGCAGAGTATTTGTTACACAGGAGCTGATGTGAATGCCAAGCAAAAGGAGGA 1860  
QY 2342 CTTATTCCTTTACATATGACAGATCTTACGGGATGTAGATGTAGCAGCTCTACTAATA 2401  
Db 1861 CTTATTCCTTTACATATGACAGATCTTACGGGATGTAGATGTAGCAGCTCTACTAATA 1920  
QY 2402 AAGTATAATGATGTCTCAATGCCAGGACAAATGGGCTTTACACCTTTTGCACGAAGCA 2461  
Db 1921 AAGTATAATGATGTCTCAATGCCAGGACAAATGGGCTTTACACCTTTTGCACGAAGCA 1980  
QY 2462 GCCCAAGGAGGACGAACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACT 2521  
Db 1981 GCCCAAGGAGGACGAACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACT 2040  
QY 2522 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTTCAGCGGATGTAGCGCT 2581  
Db 2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTTCAGCAGATGTAGTCAGCGCT 2100  
QY 2582 CTTCTGACAGAGCCATGCCCCATCTCTCTGCCCTCTTGTACAGCCCTCAAGTGTCT 2641  
Db 2101 CTTCTGACAGAGCCATGCCCCATCTCTCTGCCCTCTTGTACAGCCCTCAAGTGTCT 2160  
QY 2642 AATGTGTGAGAGCCCGAGGAGCCACTGCAGATGCTCTCTTCAGGTCCATCTAGCCCA 2701  
Db 2161 AATGTGTGAGAGCCCGAGGAGCCACTGCAGATGCTCTCTTCAGGTCCATCTAGCCCA 2220  
QY 2702 TCAAGCCCTTCTGACGACGAGCTCTTACACACTTATCTGGAGTTTTCAGAACTGTCT 2761  
Db 2221 TCAAGCCCTTCTGACGACGAGCTCTTACACACTTATCTGGAGTTTTCAGAACTGTCT 2280  
QY 2762 TCAGTAGTTAGTTCAAGTGGACAGAGGCTGCTTCCAGTTTGGAAAAAGAGGTTCCA 2821  
Db 2281 TCAGTAGTTAGTTCAAGTGGACAGAGGCTGCTTCCAGTTTGGAAAAAGAGGTTCCA 2340  
QY 2822 GGAGTAGATTTTAGCATACTCAATTCGTAGGAATCTTGGACTTTGACACCTCAATGGAT 2881  
Db 2341 GGAGTAGATTTTAGCATACTCAATTCGTAGGAATCTTGGACTTTGACACCTCAATGGAT 2400  
QY 2882 ATATTTGAGAGAACAGATCACTTTGGATGTATTTAGTTGAGATGGGCAACAAGAGCTG 2941  
Db 2401 ATATTTGAGAGAACAGATCACTTTGGATGTATTTAGTTGAGATGGGCAACAAGAGCTG 2460  
QY 2942 AAGGAGATTGGAAATCAATGCTTATGGACATAGCCCAAACTAAATTAAGGAGTTCAGAGA 3001  
Db 2461 AAGGAGATTGGAAATCAATGCTTATGGACATAGCCCAAACTAAATTAAGGAGTTCAGAGA 2520  
QY 3002 CTTATCTCGGACACAAGGCTTTAACCCATATTTAACTTTGAAACACCTCTGGTAGTGA 3061  
Db 2521 CTTATCTCGGACACAAGGCTTTAACCCATATTTAACTTTGAAACACCTCTGGTAGTGA 2580  
QY 3062 ACAATTTCTTATAGATCTGTCTCTGATGATAAGAGTTTTCAGTCTGTGGAGCAAGAGATG 3121  
Db 2581 ACAATTTCTTATAGATCTGTCTCTGATGATAAGAGTTTTCAGTCTGTGGAGCAAGAGATG 2640

QY 3122 CAAAGTACAGTTTCGAGAGCACAGAGATGAGGTTCATGCAGGTGGAATCTTCAACAGATAC 3181  
Db 2641 CAAAGTACAGTTTCGAGAGCACAGAGATGAGGTTCATGCAGGTGGAATCTTCAACAGATAC 2700  
QY 3182 AATATTCTCAAGATTCAGAAGTTTGTACAAGAACTATGGGAAAGATACACTCACCGG 3241  
Db 2701 AATATTCTCAAGATTCAGAAGTTTGTACAAGAACTATGGGAAAGATACACTCACCGG 2760  
QY 3242 AGAAAAAGATTTCTCAAGAAACACCAACCATGCCAATGAACGAATGCTATTTCATGGG 3301  
Db 2761 AGAAAAAGATTTCTCAAGAAACACCAACCATGCCAATGAACGAATGCTATTTCATGGG 2820  
QY 3302 TCTCCTTTTGTGAATCAATTTATCCACAAGGCTTTGTGAAAGGATCGGTACATAGGT 3361  
Db 2821 TCTCCTTTTGTGAATCAATTTATCCACAAGGCTTTGTGAAAGGATCGGTACATAGGT 2880  
QY 3362 GGTATGTTGGAGCTGGCATTTATTTTGTGTAAGACTCTTCCAAAAGCAATCAATATGTA 3421  
Db 2881 GGTATGTTGGAGCTGGCATTTATTTTGTGTAAGACTCTTCCAAAAGCAATCAATATGTA 2940  
QY 3422 TATGGAATTTGGAGGAGTACTGGGTGTCAGTTTCACAAAAGACAGATCTTGTACATTTGC 3481  
Db 2941 TATGGAATTTGGAGGAGTACTGGGTGTCAGTTTCACAAAAGACAGATCTTGTACATTTGC 3000  
QY 3482 CACAGGAGCTGCTCTTTTGGCGGGTAACTTGGGAAAGTCTTCTCGCAGTTTCACTGCA 3541  
Db 3001 CACAGGAGCTGCTCTTTTGGCGGGTAACTTGGGAAAGTCTTCTCGCAGTTTCACTGCA 3060  
QY 3542 ATGAAATGACATCTCTCTCCAGGTTCATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGT 3601  
Db 3061 ATGAAATGACATCTCTCTCCAGGTTCATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGT 3120  
QY 3602 GGCTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTA 3661  
Db 3121 GGCTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTA 3180  
QY 3662 ATTACTTACCAGATTTATGAGGCTGAAAGTATGTCGATGGATAAATAGTATTTTAAAGA 3721  
Db 3181 ATTACTTACCAGATTTATGAGGCTGAAAGTATGTCGATGGATAAATAGTATTTTAAAGA 3240  
QY 3722 ACTAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGGCCTCTAGGTTTACTCCTT 3781  
Db 3241 ACTAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGGCCTCTAGGTTTACTCCTT 3300  
QY 3782 TGCTGAAAAA 3794  
Db 3301 TGCTGAAAAA 3313

RESULT 11

AAF63837 standard; DNA; 3508 BP.

ID AAF63837 standard; DNA; 3508 BP.

XX AAF63837;

XX 05-APR-2001 (first entry)

DT Human tankyrase2 related coding sequence SEQ ID NO: 1.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

XX Inflammatory disorder; ds.

XX Homo sapiens.

OS WO200100849-A1.

PN 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
XX WPI: 2001-102896/11.  
DR P-PSDB: AAB66278.  
XX  
PT New tankyrase2 polypeptides, useful for treating conditions mediated by  
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
PT inflammatory and autoimmune disorders.  
XX  
PS Disclosure; Page 105-109; 242pp; English.  
XX  
CC The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.  
XX  
SQ Sequence 3508 BP; 1046 A; 710 C; 843 G; 909 T; 0 other;  
  
Query Match 86.8%; Score 3296.2; DB 22; Length 3508;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 403 AGTGTGGGCGAAGACGTAGTGAATATTTGCTTCAGATGCGCAAGTCTCCAAGC 462  
DB |||||||  
QY 463 ACCTGATGATGGGGCTTATCCTCTTCATATGATGCTCTTTTGGTCATGCTGAAGT 522  
DB |||||||  
QY 523 AGTCAATCTCTTTTGGACATGCTGCGACACCCCAATGCTCGAGATAATTTGAATATAC 582  
DB |||||||  
QY 583 TCCTCTCCATGAAGCTGCAATTAAGGAAGATTTGATGTTTTCATTTGCTTTACAGCA 642  
DB |||||||  
QY 643 TGAGCTGAGCCACCATCGAATACAGATGGAAGGACAGCATTTGATTTAGCAGATCC 702  
DB |||||||  
QY 648 TGAGCTGAGCCACCATCGAATACAGATGGAAGGACAGCATTTGATTTAGCAGATCC 507  
QY 703 ATCTGCCAAGACAGTCTTACTTGGTGAATATTAAGAAAGATGAACCTTAGAAGTGCCAG 762  
DB |||||||  
QY 763 GAGTGGCAATGAAGAAAATGATGGCTTACTACACCATTTAAATGTCACTGCCACGC 822  
DB |||||||  
QY 823 AGTGTATGCGAGAACTCAACTCCATACATTTGGCAGCAGGATATACAGATTAAGAT 882  
DB |||||||  
QY 883 TGTACAGCTGTTTACTGCAACATGGAGCTGATGCTCCATGCTAAAGATAAAGTGATCTGCT 942  
DB |||||||  
QY 943 ACCATTACAAATGCTGTTTATGGTCAATATGAAGTAACCTGAACCTTTTGGTCAAGCA 1002  
DB |||||||  
QY 1003 TGGTCCCTGTGAATGCAATGAGCTTTGGCAATTCCTCTCTATGAGGAGCTTC 1062  
DB |||||||  
QY 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTTAAGTTATGTTGCGAGACCCCAACACTGCT 1122  
DB |||||||

DB 868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTTAAAGTTATGGTGCAGACCCCAACTGCT 927  
QY 1123 CAATTGTCAATAAAAGTGTCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATT 1182  
DB |||||||  
QY 928 CAATTGTCAATAAAAGTGTCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATT 987  
QY 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTCAAGCTGCAGAGAAAGCTGATGTTACTCG 1242  
DB |||||||  
QY 988 AGCATATGAATTTAAAGGCCACTCGTTGCTCAAGCTGCAGAGAAAGCTGATGTTACTCG 1047  
QY 1243 AATCAAAAACATCTCTCTGGAATGTTCAAGCATCTCAACACATGAAC 1302  
DB |||||||  
QY 1048 AATCAAAAACATCTCTCTGGAATGTTCAAGCATCTCAACACATGAAC 1107  
QY 1303 AGCATTTGATTTGCTGCTGCTATCCATATCCCAAGAAAGCAATATGTTGAAGTGT 1362  
DB |||||||  
QY 1108 AGCATTTGATTTGCTGCTGCTATCCATATCCCAAGAAAGCAATATGTTGAAGTGT 1167  
QY 1363 GCTAAGAAAGGAGCAACATCAATGAAGCACTAAAGAAATTTCTTGACTCTCTGCGAGT 1422  
DB |||||||  
QY 1168 GCTAAGAAAGGAGCAACATCAATGAAGCACTAAAGAAATTTCTTGACTCTCTGCGAGT 1227  
QY 1423 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGTGTTGAAGCAATGAAGCAAGT 1482  
DB |||||||  
QY 1228 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGTGTTGAAGCAATGAAGCAAGT 1287  
QY 1483 TAATGCTCTGATATCTTGTGTCAGACTTCTCTACAGAGCTGCATATTTGTTGCTCATCT 1542  
DB |||||||  
QY 1288 TAATGCTCTGATATCTTGTGTCAGACTTCTCTACAGAGCTGCATATTTGTTGCTCATCT 1347  
QY 1543 ACAACCTCGCCGCTACTCTCTGAGCTATGGGTGTGATCCTTAACATTTATCCCTTCAGGG 1602  
DB |||||||  
QY 1348 ACAACCTCGCCGCTACTCTCTGAGCTATGGGTGTGATCCTTAACATTTATCCCTTCAGGG 1407  
QY 1603 CTTTACTGCTTTTACAGATGGAATGAAAATGTACAGCAACTCTCTCAAGAGGATATCTC 1662  
DB |||||||  
QY 1468 ATTAGGTAATTCAGAGCAGACACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1722  
DB |||||||  
QY 1723 AACTGTAAAAAAGCTGTACTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1782  
DB |||||||  
QY 1528 AACTGTAAAAAAGCTGTACTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCA 1587  
QY 1783 GTCTACACCACTTCAATTTTGCAGTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1842  
DB |||||||  
QY 1588 GTCTACACCACTTCAATTTTGCAGTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1647  
QY 1843 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTTGACCTTTGCACATGC 1902  
DB |||||||  
QY 1648 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTTGACCTTTGCACATGC 1707  
QY 1903 ATGTTCTTATGGACATTTGAAGTTGCGAAGCTTCTTTTAAACATGGAGCAGTAGTTAA 1962  
DB |||||||  
QY 1708 ATGTTCTTATGGACATTTGAAGTTGCGAAGCTTCTTTTAAACATGGAGCAGTAGTTAA 1767  
QY 1963 TGTAGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAGAGAAATATGA 2022  
DB |||||||  
QY 1768 TGTAGCTGATTTTGAAGATTTTACCTTTTACATGAAGCAGCAGCAAGAGAAATATGA 1827  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCACAAAAAAGAGGATGAAA 2082  
DB |||||||  
QY 1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACAAAAAAGAGGATGAAA 1887  
QY 2083 TACTCTTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 2142  
DB |||||||  
QY 1888 TACTCTTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGA 1947  
QY 2143 TGCAGCTTTCTAGATGCTGCCAAGAGGCTTTTATGAGCAGAGTGAAGAGTTGCTTTC 2202  
DB |||||||  
QY 1948 TGCAGCTTTCTAGATGCTGCCAAGAGGCTTTTATGAGCAGAGTGAAGAGTTGCTTTC 2007

```
QY 2203 TCCTGATAAATGTAATTCGCGCGATACCCAGGCGAGACATTCACACCTTTTACATTTAGC 2262
Db 2008 TCCTGATAAATGTAATTCGCGCGATACCCAGGCGAGACATTCACACCTTTTACATTTAGC 2067
QY 2263 AGCTGGTTATATAATTTAGAAAGTTGCAGAGTATTTCTTACAACACGAGCTGATGTGAA 2322
Db 2058 AGCTGGTTATATAATTTAGAAAGTTGCAGAGTATTTCTTACAACACGAGCTGATGTGAA 2127
QY 2323 TGCCCAAGACAAAAGGAGGACTTATTCCTTTACATAAATGCAGCATCTTTACGGGATGTAGA 2382
Db 2128 TGCCCAAGACAAAAGGAGGACTTATTCCTTTACATAAATGCAGCATCTTTACGGGATGTAGA 2187
QY 2383 TGTAGCAGCTCTACTATAAAGTATATGATGTGTCATGCGCAGCAAAATGGGCTTT 2442
Db 2188 TGTAGCAGCTCTACTATAAAGTATATGATGTGTCATGCGCAGCAAAATGGGCTTT 2247
QY 2443 CACACCTTTTGACGAGAGCGCCCAAGGAGGACGACACGCTTTGCTCTTTGTTGCTAGC 2502
Db 2248 CACACCTTTTGACGAGAGCGCCCAAGGAGGACGACACGCTTTGCTCTTTGTTGCTAGC 2307
QY 2503 CCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2308 CCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367
QY 2563 AGCGGATGATGTACGCGCTCTTCTGACAGCAGCCATGCCGCCATCTGCTCTGCGCCTCTTG 2622
Db 2368 AGCAGATGATGTACGCGCTCTTCTGACAGCAGCCATGCCGCCATCTGCTCTGCGCCTCTTG 2427
QY 2623 TTACAGACCTCAAGTGCTCAATGTTGTGTGAGAGCCCGAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2428 TTACAGACCTCAAGTGCTCAATGTTGTGTGAGAGCCCGAGGAGCCACTGCAGATGCTCTCTC 2487
QY 2683 TTCAGGTCCTATAGCCCATCAAGCCTTCTGACGAGCAGAGTCTTGACAACTTTATCTGG 2742
Db 2488 TTCAGGTCCTATAGCCCATCAAGCCTTCTGACGAGCAGAGTCTTGACAACTTTATCTGG 2547
QY 2743 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2802
Db 2548 GAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAACAGAGGCTGCTCCAGTTT 2607
QY 2803 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2608 GGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667
QY 2863 ACTTCAGCACCTTAATGGATATATTGTAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2668 ACTTCAGCACCTTAATGGATATATTGTAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727
QY 2923 GATGGGCACAAAGGAGCTCAAGGAGATTTGGAATCAATGCTTTATGGACATAGGCACAACT 2982
Db 2728 GATGGGCACAAAGGAGCTCAAGGAGATTTGGAATCAATGCTTTATGGACATAGGCACAACT 2787
QY 2983 AATTAAAGAGTCGAGAGACTTATCTCCGACACACAGGCTTTAACCCATATTAACTTT 3042
Db 2788 AATTAAAGAGTCGAGAGACTTATCTCCGACACACAGGCTTTAACCCATATTAACTTT 2847
QY 3043 GAACACCTCTGTGTAGTGAACAATTTCTTATAGATCTGCTCTGATGATAAGAGTTTCA 3102
Db 2848 GAACACCTCTGTGTAGTGAACAATTTCTTATAGATCTGCTCTGATGATAAGAGTTTCA 2907
QY 3103 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTTCGAGAGCAGACAGATGGAGGTCTAGCAGG 3162
Db 2908 GTCTGTGGAGGAAGAGATGCAAGGTACAGTTTCGAGAGCAGACAGATGGAGGTCTAGCAGG 2967
QY 3163 TGGAAATCTTCAACAGATACATATTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTATG 3222
Db 2968 TGGAAATCTTCAACAGATACATATTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTATG 3027
QY 3223 GGAAGATACACTCCCGGAGAAAAGAGTTTCTGAAGAAAACCAACCACTGCAATGA 3282
Db 3028 GGAAGATACACTCCCGGAGAAAAGAGTTTCTGAAGAAAACCAACCACTGCAATGA 3087
```

```
QY 3283 ACGAATGCTATTTCATGGCTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3088 ACGAATGCTATTTCATGGGTCCTCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147
QY 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTTC 3402
Db 3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTTC 3207
QY 3403 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGAGTACTGGGTGTCAGTTTCAAAAGA 3462
Db 3208 CAAAAGCAATCAATATGATATGGAATTTGGAGAGGAGTACTGGGTGTCAGTTTCAAAAGA 3267
QY 3463 CAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTCCCGGTTAACCTTGGGAAAGTC 3522
Db 3268 CAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTCCCGGTTAACCTTGGGAAAGTC 3327
QY 3523 TTTCCCTGCAAGTTTCAGTGCATGAAATGCAATTCCTCTCCAGGTCATCACTCAGTCAC 3582
Db 3328 TTTCCCTGCAAGTTTCAGTGCATGAAATGCAATTCCTCTCCAGGTCATCACTCAGTCAC 3387
QY 3583 TGTAGGCCAGGTGTAATGGCTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACA 3642
Db 3388 TGTAGGCCAGGTGTAATGGCTAGCATTTAGCTGAATATGTTTATTTACAGAGGAGAACA 3447
QY 3643 GCCTTATCTGAGTATTTAAATTTACAGATTTAGAGCCCTGAAGCTATGTCGATGG 3702
Db 3448 GCCTTATCTGAGTATTTAAATTTACAGATTTAGAGCCCTGAAGCTATGTCGATGG 3507
QY 3703 A 3703
Db 3508 A 3508

RESULT 12
AAC66823
ID AAC66823 standard; cdna; 4493 BP.
AC AAC66823;
XX
XX 27-FEB-2001 (first entry)
XX
DE Human tankyrase II coding sequence SEQ ID NO: 1.
XX
KW Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS
FT 1..4002
FT /*tag= a
FT /transl_except= (pos:595..597,aa:Xaa)
FT /transl_except= (pos:625..630,aa:XaaXaa)
FT /transl_except= (pos:637..639,aa:Xaa)
FT /transl_except= (pos:694..696,aa:Xaa)
FT /transl_except= (pos:724..726,aa:Xaa)
FT /transl_except= (pos:730..732,aa:Xaa)
FT /transl_except= (pos:763..765,aa:Xaa)
FT /transl_except= (pos:790..792,aa:Xaa)
FT /transl_except= (pos:2176..2178,aa:Xaa)
FT /transl_except= (pos:2380..2382,aa:Xaa)
FT /transl_except= (pos:2398..2400,aa:Xaa)
FT /transl_except= (pos:2464..2469,aa:XaaXaa)
FT /transl_except= (pos:3163..3165,aa:Xaa)
FT /note= "Xaa=unknown"
FT /product= "tankyrase II"
FT /partial
XX
XX WO200061813-A1.
XX
XX 19-OCT-2000.
XX
XX 10-APR-2000; 2000WO-US09558.
XX
```



```
Db 2352 TACTCTTTTGGATCTGTGTAAGATGGAGANACAGATATTCAGATNTGCTTAGGGGAGA 2411
Qy 2143 TGCAGCTTTGTAGATGCTGCAAGAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2412 TGCAGCTTTGTAGATGCTGCAAGAGAGGTTGTTTAGCCAGAGTGAAGAAGTTNTWTT 2471
Qy 2203 TCCTGATAATGTAATGTCGCCGATACCCAAAGGAGACATTCACACCTTTACATTTAGC 2262
Db 2472 TCCTGATAATGTAATGTCGCCGATACCCAAAGGAGACATTCACACCTTTACATTTAGC 2531
Qy 2263 AGCT-----GGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACAGGAGCTGA 2316
Db 2532 AGTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2591
Qy 2317 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATCCAGCATCTTACGGGCA 2376
Db 2592 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2651
Qy 2377 TGTAGATGACAGCTCTACTAATAAGTATAATGTCATGTGTCATGCCAGGACAAATG 2436
Db 2652 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2711
Qy 2437 GCCTTTCACACCTTTCGACGAGAGCGCCCAAGAGGAGCAACAGCTTTGTGCTTTGTT 2496
Db 2712 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2771
Qy 2497 GCTAGCCCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGACAAACACCTTTAGATT 2556
Db 2772 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2831
Qy 2557 AGTTTCAGCGGATGATGTCAGCGCTCTTCTGCACAGAGCCATGCCCCCATCTGCTGCC 2616
Db 2832 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2891
Qy 2617 CTCCTGTTACAAGCCTCAAGTCTCAATGGTGTGAGAAGCCAGGAGCCAGCTGCAGATGC 2676
Db 2892 CTCCTGTTACAAGCCTCAAGTCTCAATGGTGTGAGAAGCCAGGAGCCAGCTGCAGATGC 2951
Qy 2677 TCCTCTCTCAGCTCCATCTAGCCCATCAAGCCTTCTCAGCCAGCAGCTCTTGACAACTT 2736
Db 2952 TCCTCTCTCAGCTCCATCTAGCCCATCAAGCCTTCTCAGCCAGCAGCTCTTGACAACTT 3011
Qy 2737 ATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTATGTTCAAGTGGAAACAGAGGTTGCTTC 2796
Db 3012 ATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTATGTTCAAGTGGAAACAGAGGTTGCTTC 3071
Qy 2797 CAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 2856
Db 3072 CAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 3131
Qy 2857 TCCTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGATTT 2916
Db 3132 TCCTGGACTTGAGCACCTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGATTT 3191
Qy 2917 AGTTGAGATGGGGCAACAGGAGCTGAAGGAGATGGAATCAATGCTTATGGACATAGGCA 2976
Db 3192 AGTTGAGATGGGGCAACAGGAGCTGAAGGAGATGGAATCAATGCTTATGGACATAGGCA 3251
Qy 2977 CAAACTAATTAAGGAGCTCGAGAGCTTATCTCCGGACAAACAGGCTTACCCATATTT 3036
Db 3252 CAAACTAATTAAGGAGCTCGAGAGCTTATCTCCGGACAAACAGGCTTACCCATATTT 3311
Qy 3037 AACTTTGAACACCTCTGCTAGTGGAAACAATCTTATAGATCTGTCTCCTGATGATAAGA 3096
Db 3312 AACTTTGAACACCTCTGCTAGTGGAAACAATCTTATAGATCTGTCTCCTGATGATAAGA 3371
Qy 3097 GTTTCAGCTGTGGAGGAGAGATGCAAGTACAGTTCCAGAGCACAGAGATGGAGGTCA 3156
Db 3372 GTTTCAGCTGTGGAGGAGAGATGCAAGTACAGTTCCAGAGCACAGAGATGGAGGTCA 3431
Qy 3157 TGCAGGTGGAATCTTCAACAGATACAATATCTCAAGATTCAGAGGTTTCTACAAGAA 3216
Db 3432 TGCAGGTGGAATCTTCAACAGATACAATATCTCAAGATTCAGAGGTTTCTACAAGAA 3491
```

```
Qy 3217 ACTATG-----GGAAGATACACTCACCGGAGAAAAAGAGTTTC 3255
Db 3492 NNNNNNNNGAGCCAAAGATTCGGCACAGGAAAGATACACTCACCGGAGAAAGAGTTTC 3551
Qy 3256 TGAAGAAAACCAACACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTGTGAA 3315
Db 3552 TGAAGAAAACCAACACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTGTGAA 3611
Qy 3316 TGCATTTATCCCAAAAGGCTTTCATGAAAGGCATGCGTACATAGTGGTATGTTTGGAGC 3375
Db 3612 TGCATTTATCCCAAAAGGCTTTCATGAAAGGCATGCGTACATAGTGGTATGTTTGGAGC 3671
Qy 3376 TGCATTTATTTTGTGAAAACTCTTCCAAAAGCAATCAATATGATATATGGAATTCGAGG 3435
Db 3672 TGCATTTATTTTGTGAAAACTCTTCCAAAAGCAATCAATATGATATGGAATTCGAGG 3731
Qy 3436 AGTACTGGGTGTCAGTTTCACAAAAGACAGATCTTCTTACATTTGCCAGGAGCTGCT 3495
Db 3732 AGTACTGGGTGTCAGTTTCACAAAAGACAGATCTTCTTACATTTGCCAGGAGCTGCT 3791
Qy 3496 CTTTTCGCCGGGTAAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCATGAAATGGCACA 3555
Db 3792 CTTTTCGCCGGGTAAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCATGAAATGGCACA 3851
Qy 3556 TTCTCTCTCCAGGTCATCAGTCAGTCTAGGAGGAGTAAATGGCTTAGCATTAGC 3615
Db 3852 TTCTCTCTCCAGGTCATCAGTCAGTCTAGGAGGAGTAAATGGCTTAGCATTAGC 3911
Qy 3616 TGAATATGTTTATTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTTACCAGAT 3675
Db 3912 TGAATATGTTTATTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTTACCAGAT 3971
Qy 3676 TATGAGGCTTGAAGGTATGTCGATGGATAAATAGTATTTTAAAGAACTAATTCACCTG 3735
Db 3972 TATGAGGCTTGAAGGTATGTCGATGGATAAATAGTATTTTAAAGAACTAATTCACCTG 4031
Qy 3736 AACCTAAATCATCAAGCAGCAGTGGCCCTACTGCTTTACTCCTTTGCTGAAAAAAA 3794
Db 4032 AACCTAAATCATCAAGCAGCAGTGGCCCTACTGCTTTACTCCTTTGCTGAAAAAAA 4090

RESULT 13
AAC66824
ID AAC66824 standard; cDNA; 4296 BP.
XX
AC AAC66824;
XX
DT 27-FEB-2001 (first entry)
XX
Human tankyrase II coding sequence SEQ ID NO: 3.
Human; tankyrase II; telomere length; signal transduction; ss.
XX
Homo sapiens.
Key Location/Qualifiers
CDS 3..4373
/tag= a
/transl_except= (pos:182..184,aa:Xaa)
/transl_except= (pos:350..352,aa:Xaa)
/transl_except= (pos:385..387,aa:Xaa)
/transl_except= (pos:397..399,aa:Xaa)
/transl_except= (pos:434..436,aa:Xaa)
/transl_except= (pos:473..475,aa:Xaa)
/transl_except= (pos:554..556,aa:Xaa)
/transl_except= (pos:575..580,aa:XaaXaa)
/transl_except= (pos:590..592,aa:Xaa)
/transl_except= (pos:611..613,aa:Xaa)
/transl_except= (pos:620..622,aa:Xaa)
/transl_except= (pos:662..664,aa:Xaa)
/transl_except= (pos:668..673,aa:XaaXaa)
/transl_except= (pos:677..679,aa:Xaa)
```





Db	1555	 TGAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTTGGTCAGACACTCTCTACACAGAG	1614
Qy	1524	CTGCATATTCTGGTCATCTACAAACCTGCGGCTACTCCTCAGCTATTGGGTGTGATCCCTA	1583
Db	1615	CTGCATATTCTGGTCATCTACAAACCTGCGGCTACTCCTCAGCTATTGGGTGTGATCCCTA	1674
Qy	1584	ACATTATATCCCTTCAGGGCTTTACTTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC	1643
Db	1675	ACATTATATCCCTTCAGGGCTTTACTTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC	1734
Qy	1644	TCCTCCAAGAGGTATCTCATTTAGTAAATTACAGAGGCACAGACAATTTGCTGGAAGCTG	1703
Db	1735	TCCTCCAAGAGGTATCTCATTTAGTAAATTACAGAGGCACAGACAATTTGCTGGAAGCTG	1794
Qy	1704	CAAAAGCTCGAGATGTCGAAACTGTAAAAAACTCTGTACTGTTTCAGAGTGTCAACTGCA	1763
Db	1795	CAAGGCTCGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCA	1854
Qy	1764	GAGACATTTGAAGGGCTCAGTCTACACCACCTTCATTTTGCAGCTGGGTATATAACAGAGTGT	1823
Db	1855	GAGACATTTGAAGGGCTCAGTCTACACCACCTTCATTTTGCAGCTGGGTATATAACAGAGTGT	1914
Qy	1824	CCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGGAGGCC	1883
Db	1915	CCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGGAGGCC	1974
Qy	1884	TTGTACCTTTTGCACAATGCATGCTCTTATGAGACATTAAAGTTGCGAGACTCTCTGTTA	1943
Db	1975	TTGTACCTTTTGCACAATGCATGCTTATGAGACATTAAAGTTGCGAGACTCTCTGTTA	2034
Qy	1944	AACATGGAGCAGTAGTTAATGTAGCTGATTTATGAAAATTTACACCTTTACATGAAGCAG	2003
Db	2035	AACATGGAGCAGTAGTTAATGTAGCTGATTTATGAAAATTTACACCTTTACATGAAGCAG	2094
Qy	2004	CAGCAAAAGGAAAATATGAAAATTTGCAAACTTCTCCTCCAGCATGGTGCAGACCTACCA	2063
Db	2095	CAGCAAAAGGAAAATATGAAAATTTGCAAACTTCTCCTCCAGCATGGTGCAGACCTACCA	2154
Qy	2064	AAAAAACAGGGATGAAAATACTCCTTTGGATCTTGTTTAAAGATGGAGATACAGATATTC	2123
Db	2155	AAAAAACAGGAAAATCTNTTTGGATCTTGTTTAAAGATGGAGANACAGATATTC	2214
Qy	2124	AAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCA	2183
Db	2215	AAGATNTGCTTTAGGGAGATGCAGTTTGTAGATGCTGCCAAGAAGGGTTGTTTAGCCA	2274
Qy	2184	GAGTGAAGAAGTTGCTCTTCTCTGATAATGTAAATTTGCCGGATACCCAGGCGACATT	2243
Db	2275	GAGTGAAGAAGTTNTNTTTCTCTGATAATGTAAATTTGCCGNGATACCCAGGCGACATT	2334
Qy	2244	CAACACCTTTTACATTTAGCAGCTGTATATAATTTAGAGTTGCAGAGTATTTGTTAC	2303
Db	2335	CAACACCTTTTACATTTAGCA--GGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2391
Qy	2304	AACAGGAGCTGATGTAATGCCAAGACAAAAGGAGGACTTATTCCTTTACATAATCAG	2363
Db	2392	NN	2451
Qy	2364	CATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAGTATAATGTCATGCTCAATG	2423
Db	2452	NN	2511
Qy	2424	CCACGGACAAATGGGCTTTCACACCTTTGCACGAAGACGCCAAAAGGACACACAGC	2483
Db	2512	NN	2571
Qy	2484	TTTGTGCTTTGTCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGACAAA	2543
Db	2572	NN	2631
Qy	2544	CACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGCACAGACCATGCCCC	2603

[illegible]

QY 3675 TTATGAGGCGCTGAAGGTATGGTCGATGATAAATAGTATTATTTAAGAACTAATTCACCT 3734  
|||||  
Db 3772 TTATGAGGCGCTGAAGGTATGGTCGATGATAAATAGTATTATTTAAGAACTAATTCACCT 3831  
QY 3735 GAACCTAAATCATCAAGAGCAGAGTGGCGCTCTACGTTTACTCCTTTGCTGAAAAA 3794  
|||||  
Db 3832 GAACCTAAATCATCAAGAGCAGAGTGGCGCTCTACGTTTACTCCTTTGCTGAAAAA 3891  
RESULT 14  
AAC89706  
ID AAC89706 standard; cdna; 5005 BP.  
XX AC AAC89706;  
XX 08-MAR-2001 (first entry)  
XX Human adult T-cell leukaemia cDNA MO-BC-203.  
DE XX  
XX Human; cytostatic; gene therapy; vaccine; breast cancer;  
KW T cell leukaemia; cancer associated antigen; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200072021-A2.  
PN 30-NOV-2000.  
PD 24-MAY-2000; 2000WO-US14391.  
PF 26-MAY-1999; 99US-0320092.  
PR (LUDW-) LUDWIG INST CANCER RES.  
XX PA  
XX Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;  
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;  
PI Sahlin U, Piruendschuh M, Old LJ, Knuth A, Jager E;  
XX  
XX WPI; 2001-032067/04.  
DR P-PSDB; AAB48574.  
XX  
XX Diagnosing disorder characterized by expression of human cancer  
PT associated antigen precursor, involves detecting interaction of an  
PT agent with novel NA Group 1 nucleic acid molecule encoding the antigen  
PT precursor -  
XX  
XX Example 8; Page 82-84; 87pp; English.  
PS  
XX The present sequence is given in a specification relating to nucleic  
CC acids and encoded polypeptides which are cancer associated antigens  
CC expressed in patients afflicted with breast cancer and/or T cell  
CC leukaemia. A disorder characterised by expression of human cancer  
CC associated antigen may be diagnosed by contacting a biological sample  
CC with an agent that specifically binds to the nucleic acid that encodes  
CC the antigen, complexed with human leukocyte antigen molecule, and  
CC determining the interaction between the agent and the nucleic acid.  
CC Cancer associated antigens, the nucleotides encoding them, antibodies  
CC against them and the pharmaceutical compositions comprising them are  
CC useful for diagnosing, monitoring and treating the diseases characterised  
CC by the expression of one or more cancer associated antigens.  
XX  
SQ Sequence 5005 BP; 1492 A; 917 C; 1015 G; 1571 T; 10 other;  
Query Match 73.1%; Score 2774.6; DB 22; Length 5005;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2810; Conservative 0; Mismatches 4; Indels 3; Gaps 3;  
QY 978 AAGTAACCTGAACCTTTTGGTCAAGCATGGTGGCTGTGTAAATGCAATGGACTTGTGGCAAT 1037  
|||||  
Db 1 AAGTAACCTGAACCTTTTGGTCAAGCATGGTGGCTGTGTAAATGCAATGGACTTGTGGCAAT 60  
QY 1038 TCACTCCTCTTCATGAGGCGAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAA 1097  
|||||

Db 61 TCACTCCTCTTCATGAGGCGAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAA 120  
|||||  
QY 1098 GTTATGTTGAGAGCCCAACACTGCTCAATTTGTCAATAAAAGTGCATATAGACTTGGCTC 1157  
|||||  
Db 121 GTTATGTTGAGAGCCCAACACTGCTCAATTTGTCAATAAAAGTGCATATAGACTTGGCTC 180  
|||||  
QY 1158 CCACACCACAGTTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1217  
|||||  
Db 181 CCACACCACAGTTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240  
|||||  
QY 1218 CTGCACGAGAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTTGGAAATGGTGAAT 1277  
|||||  
Db 241 CTGCACGAGAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTTGGAAATGGTGAAT 300  
|||||  
QY 1278 TCAAGCATCTCAACACATGAACAGCATTTGATTTGCTGCTGCATCTCCATATCCCA 1337  
|||||  
Db 301 TCAAGCATCTCAACACATGAACAGCATTTGATTTGCTGCTGCATCTCCATATCCCA 360  
|||||  
QY 1338 AAGAAAAGCAATATGTGAAGTGTGCTAAGAAAAGAGCAAAACATCAATGAAAAGACTA 1397  
|||||  
Db 361 AAGAAAAGCAATATGTGAAGTGTGCTAAGAAAAGAGCAAAACATCAATGAAAAGACTA 420  
|||||  
QY 1398 AAGAAATCTTGAACCTCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1457  
|||||  
Db 421 AAGAAATCTTGAACCTCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480  
|||||  
QY 1458 TAGTGGTGAACATGAAGCAAAAGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1517  
|||||  
Db 481 TAGTGGTGAACATGAAGCAAAAGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540  
|||||  
QY 1518 ACAGAGCTGCATATTTGGTGCATCTACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTG 1577  
|||||  
Db 541 ACAGAGCTGCATATTTGGTGCATCTACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTG 600  
|||||  
QY 1578 ATCTTAACATATATCCCTTTCAGGGCTTTTACTGCTTTACAGATGGGAAATGAAATGTAC 1637  
|||||  
Db 601 ATCTTAACATATATCCCTTTCAGGGCTTTTACTGCTTTACAGATGGGAAATGAAATGTAC 660  
|||||  
QY 1638 AGCAACTCCTCCAAAGAGGGTATCTCATTAGTAAATTCAGAGGCAGACAGCAATGGCTGG 1697  
|||||  
Db 661 AGCAACTCCTCCAAAGAGGGTATCTCATTAGTAAATTCAGAGGCAGACAGCAATGGCTGG 720  
|||||  
QY 1698 AAGTGCAAAAGCTGGAGATGTCGAAACTGTGAAAAAAGCTGTGTTACAGAGTGTCA 1757  
|||||  
Db 721 AAGTGCAAAAGCTGGAGATGTCGAAACTGTGAAAAAAGCTGTGTTACAGAGTGTCA 780  
|||||  
QY 1758 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACACTTTCATTTTGCAGCTGGGTATAA 1817  
|||||  
Db 781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACACTTTCATTTTGCAGCTGGGTATAA 840  
|||||  
QY 1818 GAGTGTCCGGTGGGAATATCTGCTACAGCATGAGCTGATGTCATGCTTAAAGATAAAG 1877  
|||||  
Db 841 GAGTGTCCGGTGGGAATATCTGCTACAGCATGAGCTGATGTCATGCTTAAAGATAAAG 900  
|||||  
QY 1878 GAGGCTCTGTACCTTTGCAACAATGCATGTTCTTATGACATTTATGAAGTTGCAAACTTC 1937  
|||||  
Db 901 GAGGCTCTGTACCTTTGCAACAATGCATGTTCTTATGACATTTATGAATTTGCAAACTTC 960  
|||||  
QY 1938 TTGTTTAAACATGAGCAGTATGTTAAATGATGCTGATTTATGAAATTTTACACCTTTTACATG 1997  
|||||  
Db 961 TTGTTTAAACATGAGCAGTATGTTAAATGATGCTGATTTATGAAATTTTACACCTTTTACATG 1020  
|||||  
QY 1998 AAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2057  
|||||  
Db 1021 AAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 1080  
|||||  
QY 2058 CTACCAAAAAAAGAGGATGGAATTAAGTCTTGGATCTTGTAAAGATGGAGATACAG 2117  
|||||  
Db 1081 CTACCAAAAAAAGAGGATGGAATTAAGTCTTGGATCTTGTAAAGATGGAGATACAG 1140  
|||||  
QY 2118 ATATTCAAGATCTGTTAGGGGAGATGCAGCTTGTGATGATGCTGCAAGAGGGTTGTT 2177  
|||||

Db 1141 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 1200  
Qy 2178 TAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATGTAAATTTGCCGGGATACCCAAAGCA 2237  
Db 1201 TAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATGTAAATTTGCCGGGATACCCAAAGCA 1260  
Qy 2238 GACATTCACACCTTTACATTTAGCAGCTGCTTATATAATTTAGAAAGTTGCAGAGTATT 2297  
Db 1261 GACATTCACACCTTTACATTTAGCAGCTGCTTATATAATTTAGAAAGTTGCAGAGTATT 1320  
Qy 2298 TGTTCACACGGAGCTGATGTAATGCCAAGACAAAGAGGAGCTTATTCCTTTACATA 2357  
Db 1321 TGTTCACACGGAGCTGATGTAATGCCAAGACAAAGAGGAGCTTATTCCTTTACATA 1380  
Qy 2358 ATGCAGCATCTTACGGGATGATGATGAGCAGCTCTACTAATAAGATATAATGCATGTG 2417  
Db 1381 ATGCAGCATCTTACGGGATGATGATGAGCAGCTCTACTAATAAGATATAATGCATGTG 1440  
Qy 2418 TCAATGCCACGGACAAATGGGCTTTTCACACCTTTGCAGGAAGCAGCCCAAAAGGAGCA 2477  
Db 1441 TCAATGCCACGGACAAATGGGCTTTTCACACCTTTGCAGGAAGCAGCCCAAAAGGAGCA 1500  
Qy 2478 CACAGCTTTGTGCTTTGTAGTCCCATGAGCTGACCCGAGCTCTTAAAAATCAGGAAG 2537  
Db 1501 CACAGCTTTGTGCTTTGTAGTCCCATGAGCTGACCCGAGCTCTTAAAAATCAGGAAG 1560  
Qy 2538 GACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTCTGACAGCAGCA 2597  
Db 1561 GACAAACACCTTTAGATTTAGTTTCAGCGATGATGTCAGCGCTCTCTGACAGCAGCA 1620  
Qy 2598 TGCCCCCATCTCTGCGCTCTTCTTACAAGCCCTCAAGTGTCAATGGTGTGAGAAGCC 2657  
Db 1621 TGCCCCCATCTCTGCGCTCTTCTTACAAGCCCTCAAGTGTCAATGGTGTGAGAAGCC 1680  
Qy 2658 CAGGAGCCACTGCAGATGCTCTCTTCAGGTCCATCTAGCCCATCAAGCCCTTCTGCGAG 2717  
Db 1681 CAGGAGCCACTGCAGATGCTCTCT-TTCAGGTCCATCTAGCCCATCAAGCCCTTCTGCGAG 1739  
Qy 2718 CCAGCAGCTCTCAGCAACTTATCTGGAGTTTTCAGAACTGCTCTCAGTAGTTAGTTCAA 2777  
Db 1740 CCAGCAGCTCTCAGCAAC-TATCTGGAG--TTTTCAGAACTGCTCTCAGTAGTTAGTTCAA 1797  
Qy 2778 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAGAGGTTTCCAGGAGTAGATTTTAGCA 2837  
Db 1798 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAGAGGTTTCCAGGAGTAGATTTTAGCA 1857  
Qy 2838 TAACCTAATTCGTAAGGAATCTTGACCTTGACACCTTAATGGATATATTTGAGAGAGAAC 2897  
Db 1858 TAACCTAATTCGTAAGGAATCTTGACCTTGACACCTTAATGGATATATTTGAGAGAGAAC 1917  
Qy 2898 AGATCACTTTGGATGTATTAGTTGAGATGGGCAAGGAGCTGAAGGAGATTGGAATCA 2957  
Db 1918 AGATCACTTTGGATGTATTAGTTGAGATGGGCAAGGAGCTGAAGGAGATTGGAATCA 1977  
Qy 2958 ATGCTTATGGACATAGGCACAACTAATTAAGGAGTGGAGAGCTTATCTCCGGACAAAC 3017  
Db 1978 ATGCTTATGGACATAGGCACAACTAATTAAGGAGTGGAGAGCTTATCTCCGGACAAAC 2037  
Qy 3018 AAGGCTTTAACCATATTAACTTTCAACACCTCTGGTAGTGGAACTTCTTATAGATC 3077  
Db 2038 AAGGCTTTAACCATATTAACTTTGAACACCTCTGGTAGTGGAACTTCTTATAGATC 2097  
Qy 3078 TGTCTCTGATGATAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCCGAG 3137  
Db 2098 TGTCTCTGATGATAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCCGAG 2157  
Qy 3138 AGCACAGATGAGGTCTATGAGTGGAACTTCAACAGATACAAATATCTCAAGATTC 3197  
Db 2158 AGCACAGATGAGGTCTATGAGTGGAACTTCAACAGATACAAATATCTCAAGATTC 2217  
Qy 3198 AGAAGGTTTGTACACAAAGAACTATGGAAAGATACACTCACCGGAGAAAGATTCTG 3257  
Db 2218 AGAAGGTTTGTACACAAAGAACTATGGAAAGATACACTCACCGGAGAAAGATTCTG 2277

Qy 3258 AAAAAACACCAACCATGCCAATGAACGAATGCTATTTTCATGGTCTCCTTTTGTGAATG 3317  
Db 2278 AAAAAACACCAACCATGCCAATGAACGAATGCTATTTTCATGGTCTCCTTTTGTGAATG 2337  
Qy 3318 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGTGTATGTTTGGAGCTG 3377  
Db 2338 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGTGTATGTTTGGAGCTG 2397  
Qy 3378 GCATTTATTTTGTGTAATACTTCCAAAAGCAATCAATATGATATGTAATTTGGAGGAG 3437  
Db 2398 GCATTTATTTTGTGTAATACTTCCAAAAGCAATCAATATGATATGTAATTTGGAGGAG 2457  
Qy 3438 GTACTGGTCTCCAGTTCACAAAGACAGATCTTGTTCACATTTGCCACAGCAGCTGCTCT 3497  
Db 2458 GTACTGGTCTCCAGTTCACAAAGACAGATCTTGTTCACATTTGCCACAGCAGCTGCTCT 2517  
Qy 3498 TTTGCCGGGTAACTTTGGGAAAGTCTTTCTCGAGTTCAGTCAATGAAATGGCAATT 3557  
Db 2518 TTTGCCGGGTAACTTTGGGAAAGTCTTTCTCGAGTTCAGTCAATGAAATGGCAATT 2577  
Qy 3558 CTCTCCAGTCTACTCAGTCACTGCTAGTGGCCAGTGTAAATGGCTAGCAATTAGCTG 3617  
Db 2578 CTCTCCAGTCTACTCAGTCACTGCTAGTGGCCAGTGTAAATGGCTAGCAATTAGCTG 2637  
Qy 3618 AATATGTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTAAATTACTTACCAGATTA 3677  
Db 2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCTCGAGTATTTAAATTACTTACCAGATTA 2697  
Qy 3678 TGAGGCTCGAAGGTATGGTGCATGGATTAATAGTATTTTAAAGAACTAATTCCTACTGAA 3737  
Db 2698 TGAGGCTCGAAGGTATGGTGCATGGATTAATAGTATTTTAAAGAACTAATTCCTACTGAA 2757  
Qy 3738 CCTAAATCATCAAGCAGCAGTGGCTCTACGTTTACTCCTTTGCTGAAAAAAA 3794  
Db 2758 CCTAAATCATCAAGCAGCAGTGGCTCTACGTTTACTCCTTTGCTGAAAAAAA 2814

## RESULT 15

AAF63919  
ID: AAF63919 standard; DNA; 2971 BP.

AC AAF63919;

XX  
XX 05-APR-2001 (first entry)

XX Human tankyrase2 clone consensus SEQ ID NO: 92.

XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
XX inflammatory disorder; ds.

OS Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-102896/11.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
XX inflammatory and autoimmune disorders -  
XX Example 1; Page 152-153; 242pp; English.

CC The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.  
XX  
SQ

Sequence 2971 BP; 915 A; 555 C; 642 G; 859 T; 0 other;

Query Match 64.3%; Score 2439.8; DB 22; Length 2971;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1352	TGTCACCTGTTGCTAACAAAGGAGCAACATCAATCAAAAGACTAAAGAAATTCCTGACT	1411
DB	1	TGTGAACCTGTTGCTAACAAAGGAGCAACATCAATCAAAAGACTAAAGAAATTCCTGACT	60
QY	1412	CCTCGACCTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACAT	1471
DB	61	CCTCGACCTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACAT	120
QY	1472	GAAGCAAGCTTAATGCTCTGGATAATCTTGCTCAGACTTCTCTACACAGAGCTGCATAT	1531
DB	121	GAAGCAAGCTTAATGCTCTGGATAATCTTGCTCAGACTTCTCTACACAGAGCTGCATAT	180
QY	1532	TGTGGTCATCTACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTCATCCTTAACATTATA	1591
DB	181	TGTGGTCATCTACAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTCATCCTTAACATTATA	240
QY	1592	TCCCTTCAGGGCTTTACTGTTTACAGATGGGAAATGAAATGTACAGAACTCCCTCCAA	1651
DB	241	TCCCTTCAGGGCTTTACTGTTTACAGATGGGAAATGAAATGTACAGAACTCCCTCCAA	300
QY	1652	GAGGTATCTCATTAGTAAATTCAGAGGAGACAGACAATTCCTGGAAGCTGCAAGGCT	1711
DB	301	GAGGTATCTCATTAGTAAATTCAGAGGAGACAGACAATTCCTGGAAGCTGCAAGGCT	360
QY	1712	GGAGATGTCGAAACTGTAAAAAAGCTGTGCTACTGTTTCAGAGTGTCAACTGCAGAGACATT	1771
DB	361	GGAGATGTCGAAACTGTAAAAAAGCTGTGCTACTGTTTCAGAGTGTCAACTGCAGAGACATT	420
QY	1772	GAAGGGCTCAGTCTACACCCTTCAATTTTCAGCTGGGTATTAACAGAGTGTCCGTGGTG	1831
DB	421	GAAGGGCTCAGTCTACACCCTTCAATTTTCAGCTGGGTATTAACAGAGTGTCCGTGGTG	480
QY	1832	GAATATCTGCTACAGCATGGAGCTGATGTGCTATAAGATAAAGAGGCGCTTGTACCT	1891
DB	481	GAATATCTGCTACAGCATGGAGCTGATGTGCTATAAGATAAAGAGGCGCTTGTACCT	540
QY	1892	TTGCACAAATGCATGTTCTTATGGACATTAATGAAGTTGCAGAACTTCTTGTAAACATGGA	1951
DB	541	TTGCACAAATGCATGTTCTTATGGACATTAATGAAGTTGCAGAACTTCTTGTAAACATGGA	600
QY	1952	GCAGTAGTAAATGCTAGCTGATTTATGGAATTTACACCTTTACATGAAGCAGCAGCAAAA	2011
DB	601	GCAGTAGTAAATGCTAGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAA	660
QY	2012	GGAAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTACCAAAAAAACC	2071
DB	661	GGAAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACAAAAAACC	720
QY	2072	AGGGATGGAATATCTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAAGATCTG	2131
DB	721	AGGGATGGAATATCTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAAGATCTG	780
QY	2132	CTTAGGGAGATGTCAGCTTGTGCTAGATCTGCTCCAGAGAGGTTGTTTACAGAGTGAAG	2191
DB	781	CTTAGGGAGATGTCAGCTTGTGCTAGATCTGCTCCAGAGAGGTTGTTTACAGAGTGAAG	840
QY	2192	AAGTTGCTCTCTCCTGATAATGTAAATGCCCGCATACCCAGGAGCAGACATTCAACACCT	2251
DB			

DB	841	AAGTTGCTCTCTCTCTGATAATGTAATTCGCCGATACCCAGGAGCAGACATTCAACACCT	900
QY	2252	TTACATTTAGCAGCTGGTTTATAATTAATTTAGAAAGTTGCAGAGTATTTGTTAAACACGGA	2311
DB	901	TTACATTTAGCAGCTGGTTTATAATTAATTTAGAAAGTTGCAGAGTATTTGTTAAACACGGA	960
QY	2312	GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATTAATGCAGCATCTTAC	2371
DB	961	GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATTAATGCAGCATCTTAC	1020
QY	2372	GGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGTCATGTGTCATGCCACGGAC	2431
DB	1021	GGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGTCATGTGTCATGCCACGGAC	1080
QY	2432	AAATGGGCTTTCACACCTTTGACGAAAGAGCCAAAAGGAGCAACACAGCTTTTGCT	2491
DB	1081	AAATGGGCTTTCACACCTTTGACGAAAGAGCCAAAAGGAGCAACACAGCTTTTGCT	1140
QY	2492	TTGTTGCTAGCCATGGAGCTGACCCGACTCTTAAAAATCAGGAGGACAAACACCTTTA	2551
DB	1141	TTGTTGCTAGCCATGGAGCTGACCCGACTCTTAAAAATCAGGAGGACAAACACCTTTA	1200
QY	2552	GATTTAGTTTTCAGCGGATGATGTCAGCGCTCTCTGACAGCAGCATGCCCCCTCTGCT	2611
DB	1201	GATTTAGTTTTCAGCGGATGATGTCAGCGCTCTCTGACAGCAGCATGCCCCCTCTGCT	1260
QY	2612	CTGCCCTCTTGTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAAGCCAGGAGCAGCTGCA	2671
DB	1261	CTGCCCTCTTGTACAAGCCTCAAGTGTCTCAATGGTGTGAGAAAGCCAGGAGCAGCTGCA	1320
QY	2672	GATGCTCTCTCTCAGGTCCTATGAGCCATCAAGCCCTTCTGACAGCAGCAGCTTTGAC	2731
DB	1321	GATGCTCTCTCTCAGGTCCTATGAGCCATCAAGCCCTTCTGACAGCAGCAGCTTTGAC	1380
QY	2732	AACCTATCTGGGAGTTTTCAGAACTGTCCTTCAAGTAGTATAGTCAAGTGGAAACAGAGGT	2791
DB	1381	AACCTATCTGGGAGTTTTCAGAACTGTCCTTCAAGTAGTATAGTCAAGTGGAAACAGAGGT	1440
QY	2792	GCTTCCAGTTTGGAGAAAAGGAGTTCCAGAGTACAGATTTAGCATTAACCTCAATTCGTA	2851
DB	1441	GCTTCCAGTTTGGAGAAAAGGAGTTCCAGAGTACAGATTTAGCATTAACCTCAATTCGTA	1500
QY	2852	AGGAATCTTGGACTTGGACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT	2911
DB	1501	AGGAATCTTGGACTTGGACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT	1560
QY	2912	GTATTAGTTGAGATGGGCAACAGGAGCTGAAGAGATTTGAATCAATGCTTTATGGACAT	2971
DB	1561	GTATTAGTTGAGATGGGCAACAGGAGCTGAAGAGATTTGAATCAATGCTTTATGGACAT	1620
QY	2972	AGGCACAACTTAATTAAGGAGTCGAGAGCTTATCTCCGACAAACAGGCTTTAAACCA	3031
DB	1621	AGGCACAACTTAATTAAGGAGTCGAGAGCTTATCTCCGACAAACAGGCTTTAAACCA	1680
QY	3032	TATTTAACTTTGAACACCTCTGGTAGTGGAAACAAATTTATAGATCTGCTCTCTGATGAT	3091
DB	1681	TATTTAACTTTGAACACCTCTGGTAGTGGAAACAAATTTATAGATCTGCTCTCTGATGAT	1740
QY	3092	AAAGAGTTTCAGCTCTGGAGGAGAGATGCAAGATACAGTTCGAGAGCAGACAGATGGA	3151
DB	1741	AAAGAGTTTCAGCTCTGGAGGAGAGATGCAAGATACAGTTCGAGAGCAGACAGATGGA	1800
QY	3152	GGTCATGAGTGGAACTTCTCAACAGATACAAATTTCTCAAGATTCAGAAGTTTGTAAAC	3211
DB	1801	GGTCATGAGTGGAACTTCTCAACAGATACAAATTTCTCAAGATTCAGAAGTTTGTAAAC	1860
QY	3212	AAGAACTATGGGAAAGATACCTCAGCAGGAGAAAAGAGTTTCTGAAGAAAACCAAC	3271
DB	1861	AAGAACTATGGGAAAGATACCTCAGCAGGAGAAAAGAGTTTCTGAAGAAAACCAAC	1920
QY	3272	CATCCCAATGAAGAAATGCTTATTTCAATGGGCTCTCTTTTGTGAATGCAATATCCACAAA	3331
DB	1921	CATCCCAATGAAGAAATGCTTATTTCAATGGGCTCTCTTTTGTGAATGCAATATCCACAAA	1980

Qy 3332 GGCTTTGATGAAAGGCATGCGTACATAGTGGTATGTTGGAGCTGGCAATTTATTTGCT 3391  
Db 1981 GGCTTTGATGAAAGGCATGCGTACATAGTGGTATGTTGGAGCTGGCAATTTATTTGCT 2040  
Qy 3392 GAAACTCTTCCAAAGCAATCAATATGATATGGAATGGAGGAGGTACTGGGTGTCCA 3451  
Db 2041 GAAACTCTTCCAAAGCAATCAATATGATATGGAATGGAGGAGGTACTGGGTGTCCA 2100  
Qy 3452 GTTCACAAAGACAGATCTTCTTACATTTGCCACAGGAGCTGCTCTTTTGGCCGGGTAACC 3511  
Db 2101 GTTCACAAAGACAGATCTTCTTACATTTGCCACAGGAGCTGCTCTTTTGGCCGGGTAACC 2160  
Qy 3512 TTGGGAAAGTCTTTCCCTGCAAGTTCAGTGCATGCAATGAAATGGCACATTTCTCCCTCAGGTCAAT 3571  
Db 2161 TTGGGAAAGTCTTTCCCTGCAAGTTCAGTGCATGCAATGAAATGGCACATTTCTCCCTCAGGTCAAT 2220  
Qy 3572 CACTCAGTCACTGGTAGGCCAGTGTAAATGGCTAGCATTAGCTGAATATGTTATTAC 3631  
Db 2221 CACTCAGTCACTGGTAGGCCAGTGTAAATGGCTAGCATTAGCTGAATATGTTATTAC 2280  
Qy 3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTTACCAGATTTATGAGGCCTGAAGGT 3691  
Db 2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTTACCAGATTTATGAGGCCTGAAGGT 2340  
Qy 3692 ATGTCGATGGATAAATAGTATTTTAAAGAACTAATTCCTGAACTGAACTAAATCATCAA 3751  
Db 2341 ATGTCGATGGATAAATAGTATTTTAAAGAACTAATTCCTGAACTGAACTAAATCATCAA 2400  
Qy 3752 AGCAGCAGTGGCCCTACGCTTTTACTCTTTTGTGTAAGAAAAA 3794  
Db 2401 AGCAGCAGTGGCCCTACGCTTTTACTCTTTTGTGTAAGAAAAA 2443

Search completed: February 11, 2003, 13:10:06  
Job time : 752.206 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:17:58 ; Search time 115,212 Seconds  
(without alignments)  
10107.057 Million cell updates/sec

Title: US-09-843-159B-1  
Perfect score: 3797  
Sequence: 1 ctttgaagacactggatttc.....ccttctgctgaaaaa 3797

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3382.8	89.1	4512	4	US-09-350-982C-3
2	3291.8	86.7	3498	4	US-09-350-982C-4
3	1813.4	47.8	4134	4	US-09-196-387-1
4	1449.6	38.2	4491	4	US-09-196-387-7
5	1297.8	34.2	4657	4	US-09-196-387-9
6	121.2	3.2	523	4	US-09-196-387-12
7	93	2.4	5235	1	US-09-031-485-35
8	93	2.4	5235	1	US-09-031-485-36
9	93	2.4	5235	1	US-08-847-429A-35
10	93	2.4	5235	1	US-08-847-429A-36
11	93	2.4	5235	3	US-09-065-474-35
12	93	2.4	5235	3	US-09-065-474-36
13	93	2.4	5235	4	US-09-557-034-35
14	93	2.4	5235	4	US-09-557-034-36
15	93	2.4	5503	1	US-09-031-485-32
16	93	2.4	5503	1	US-09-031-485-33
17	93	2.4	5503	1	US-08-847-429A-32
18	93	2.4	5503	1	US-08-847-429A-34
19	93	2.4	5503	3	US-09-065-474-32
20	93	2.4	5503	3	US-09-065-474-33
21	93	2.4	5503	4	US-09-557-034-32
22	93	2.4	5503	4	US-09-557-034-34
23	83.4	2.2	1056	3	US-09-065-474-138
24	83.4	2.2	1056	4	US-09-065-474-140
25	83.4	2.2	1056	4	US-09-557-034-138
26	83.4	2.2	1056	4	US-09-557-034-140
27	78.8	2.1	3454	4	US-09-082-059-1

28	74.2	2.0	2505	4	US-09-291-839-3	Sequence 3, Appli
29	74.2	2.0	3025	4	US-09-291-839-1	Sequence 1, Appli
30	73.8	1.9	909	1	US-09-031-485-25	Sequence 25, Appl
c 31	73.8	1.9	909	1	US-09-031-485-26	Sequence 26, Appl
c 32	73.8	1.9	909	1	US-08-847-429A-25	Sequence 25, Appl
c 33	73.8	1.9	909	1	US-08-847-429A-26	Sequence 26, Appl
c 34	73.8	1.9	909	3	US-09-065-474-25	Sequence 25, Appl
c 35	73.8	1.9	909	3	US-09-065-474-26	Sequence 26, Appl
c 36	73.8	1.9	909	4	US-09-557-034-25	Sequence 25, Appl
c 37	73.8	1.9	909	4	US-09-557-034-26	Sequence 26, Appl
c 38	73.8	1.9	911	1	US-09-031-485-22	Sequence 22, Appl
c 39	73.8	1.9	911	1	US-09-031-485-24	Sequence 24, Appl
c 40	73.8	1.9	911	1	US-08-847-429A-22	Sequence 22, Appl
c 41	73.8	1.9	911	1	US-08-847-429A-24	Sequence 24, Appl
c 42	73.8	1.9	911	3	US-09-065-474-22	Sequence 22, Appl
c 43	73.8	1.9	911	3	US-09-065-474-24	Sequence 24, Appl
c 44	73.8	1.9	911	4	US-09-557-034-22	Sequence 22, Appl
c 45	73.8	1.9	911	4	US-09-557-034-24	Sequence 24, Appl

RESULT 1  
US-09-350-982C-3  
; Sequence 3, Application US/09350982C  
; Patent No. 6455290  
; GENERAL INFORMATION:  
; APPLICANT: Berthelsen, Jens  
; APPLICANT: Toma, Salvatore  
; APPLICANT: Isacchi, Antonella  
; TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: PHRM-0043  
; CURRENT APPLICATION NUMBER: US/09/350,982C  
; CURRENT FILING DATE: 1999-07-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 4512  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: PCR Primers  
; NAME/KEY: misc\_feature  
; LOCATION: (1124)..(1124)  
; OTHER INFORMATION: n is any nucleic acid  
; NAME/KEY: misc\_feature  
; LOCATION: (2672)..(2672)  
; OTHER INFORMATION: n is any nucleic acid  
US-09-350-982C-3

Query Match 89.1%; Score 3382.8; DB 4; Length 4512;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3381; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY	403	AGGTTTGGCGGAAGACACCTAGTTGAATATTTGCTTCAGATGGTGCAAGTGTCCAAGC	462
Db	220	AGGTTTGGCGGAAGACACCTAGTTGAATATTTGCTTCAGATGGTGCAAGTGTCCAAGC	279
QY	463	ACGTGATGATGGGGGCTTATCTCTTCATATGATGATGCTCTTTGGTCATGCTCAAGT	522
Db	280	ACGTGATGATGGGGGCTTATCTCTTCATATGATGATGCTCTTTGGTCATGCTCAAGT	339
QY	523	AGTCAATCTCTTTTGGACATGGTGACAGCCCAATGCTCGAGATAATTGGAATTATAC	582
Db	340	AGTCAATCTCTTTTGGACATGGTGACAGCCCAATGCTCGAGATAATTGGAATTATAC	399
QY	583	TCCTCTCCATGAAGCTGCAATTAAGAAAGATTGATGTTTGCATTGCTGTGTACAGCA	642
Db	400	TCCTCTCCATGAAGCTGCAATTAAGAAAGATTGATGTTTGCATTGCTGTGTACAGCA	459
QY	643	TGAGCTGAGCCCAACCATCCGAAATACAGATGAAGACAGCATTGGATTAGCAGATCC	702

ALIGNMENTS

|||||  
Db 460 TGGAGCTGAGCAACCATCCGAAATACAGATGAAGGACAGCATTTGGATTTAGCAGATCC 519  
QY 703 ATCTGCCAAAGCAGTCTTACTGGTGAATATAAGAAAGATGAACCTTPTAGAAGTGCACG 762  
Db 520 ATCTGCCAAAGCAGTCTTACTGGTGAATATAAGAAAGATGAACCTTPTAGAAGTGCACG 579  
QY 763 GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 822  
Db 580 GAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTTAAATGTCAACTGCCACGC 639  
QY 823 AAGTATGCGCAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882  
Db 640 AAGTATGCGCAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 699  
QY 883 TGTACAGCTGTTTACTGCAACATGGAGCTCATCTCCATGCTAAAGATAAAGGTGATCTGCT 942  
Db 700 TGTACAGCTGTTTACTGCAACATGGAGCTGATGTCATGCTAAAGATAAAGGTGATCTGCT 759  
QY 943 ACCATTACAAATGGCTGTTCTTATGGTCATTATGAAGTAACCTGAACCTTTTGGTCAAGCA 1002  
Db 760 ACCATTACAAATGGCTGTTCTTATGGTCATTATGAAGTAACCTGAACCTTTTGGTCAAGCA 819  
QY 1003 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACICTCTTCTATGAGGCGAGCTTC 1062  
Db 820 TGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACICTCTTCTATGAGGCGAGCTTC 879  
QY 1063 TAAGAACAGGGTGAAGTATGTCTTCTTCTTAAGTTATGTGTGCAGACCCCAACACTGCT 1122  
Db 880 TAAGAACAGGGTGAAGTATGTCTTCTTCTTAAGTTATGTGTGCAGACCCCAACACTGCT 939  
QY 1123 CAATTGTCAATAAAAGTGCTATAGACTTGGCTCCACACACAGTAAAAAGAAAGATT 1182  
Db 940 CAATTGTCAATAAAAGTGCTATAGACTTGGCTCCACACACAGTAAAAAGAAAGATT 999  
QY 1183 AGCATATGAATTTAAGGCCACTGTTGCTGCAAGCTGCACGAGAAGCTGATTACTCG 1242  
Db 1000 AGCATATGAATTTAAGGCCACTGTTGCTGCAAGCTGCACGAGAAGCTGATTACTCG 1059  
QY 1243 AATCAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCTCTCAAAACACATGAAC 1302  
Db 1060 AATCAAAAACATCTCTCTGGAATGGTGAATTTCAAGCATCTCTCAAAACACATGAAC 1119  
QY 1303 AGCATTTGATTTGCTGTGATCTCCATATCCCAAGAAAGCAAAATGATGAACTGTT 1362  
Db 1120 AGCATTTGATTTGCTGTGATCTCCATATCCCAAGAAAGCAAAATGATGAACTGTT 1179  
QY 1363 GCTAAGAAAGGAGCAACATCAATGAAGACTAAAGAAATTTGACTCTCTGCACGT 1422  
Db 1180 GCTAAGAAAGGAGCAACATCAATGAAGAACTAAAGAAATTTGACTCTCTGCACGT 1239  
QY 1423 GGCATCTGAGAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT 1482  
Db 1240 GGCATCTGAGAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT 1299  
QY 1483 TAATGCTCTGGATAATCTTGGTGTGATCTCTACACAGAGCTGCATATGTTGGTCACT 1542  
Db 1300 TAATGCTCTGGATAATCTTGGTGTGATCTCTACACAGAGCTGCATATGTTGGTCACT 1359  
QY 1543 ACAAACTCTCCGCTACTCTCAGCTATGGGTGATCTCTACATATATCCCTTCAGG 1602  
Db 1360 ACAAACTCTCCGCTACTCTCAGCTATGGGTGATCTCTACATATATCCCTTCAGG 1419  
QY 1603 CTTTACTGTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCTC 1662  
Db 1420 CTTTACTGTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCTC 1479  
QY 1663 ATTAGGTAATTCAGAGGCGACAGACAATGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1722  
Db 1480 ATTAGGTAATTCAGAGGCGACAGACAATGCTGGAAGCTGCAAGGCTGGAGATGTCGA 1539  
QY 1723 AACTGTAAAAAATCTGTACTGTTTCAGAGTGTCAACTGCACAGACATTTGAAGGGGTCA 1782  
|||||

Db 1540 AACTGTAAAAAATCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCTCA 1599  
QY 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT 1842  
Db 1600 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCT 1659  
QY 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTGTACCTTTGCACAATGC 1902  
Db 1660 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1719  
QY 1903 ATGTTCTTATGACATTTAAGATTGCAACAACCTTCTTTAAACATGGAGCAGTAGTTAA 1962  
Db 1720 ATGTTCTTATGACATTTAAGATTGCAACAACCTTCTTTAAACATGGAGCAGTAGTTAA 1779  
QY 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAAGGAAATATGA 2022  
Db 1780 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAAGGAAATATGA 1839  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGCCCTACCAAAAAACAGGAGTGA 2082  
Db 1840 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGCCCTACCAAAAAACAGGAGTGA 1899  
QY 2083 TACTCCTTTGGATCTTGTAAAGATGGATACAGATATCAAGATCTCTTAGGGGAGA 2142  
Db 1900 TACTCCTTTGGATCTTGTAAAGATGGATACAGATATCAAGATCTCTTAGGGGAGA 1959  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTC 2202  
Db 1960 TGCAGCTTTGCTAGATGCTGCCAAGAGAGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTC 2019  
QY 2203 TCCTGATAATGTAATTTGCCGCGATACCAAGCAGACATTCACACCTTTACATTTAGC 2262  
Db 2020 TCCTGATAATGTAATTTGCCGCGATACCAAGCAGACATTCACACCTTTACATTTAGC 2079  
QY 2263 AGCTGGTATATAATTTAGAAAGTTCAGAGTATTTGTTTACACAGGAGCTGATGTGA 2322  
Db 2080 AGCTGGTATATAATTTAGAAAGTTCAGAGTATTTGTTTACACAGGAGCTGATGTGA 2139  
QY 2323 TGCCCAAGACAAAGAGGAGCTTATTCCTTTACATATATGACAGCATCTTACGGGCATGTAGA 2382  
Db 2140 TGCCCAAGACAAAGAGGAGCTTATTCCTTTACATATATGACAGCATCTTACGGGCATGTAGA 2199  
QY 2383 TGTAGAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGCAAAATGGGCTTT 2442  
Db 2200 TGTAGAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGCAAAATGGGCTTT 2259  
QY 2443 CACACCTTTGCAGAGCAGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGTGCTAGC 2502  
Db 2260 CACACCTTTGCAGAGCAGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGTGCTAGC 2319  
QY 2503 CCATGGAGCTGACCGACTCTTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562  
Db 2320 CCATGGAGCTGACCGACTCTTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2379  
QY 2563 AGCGATGATGTACGGCTCTCTGACAGCAGCATGCCCATGCCCCATCTGCTCTGCCCTCTTG 2622  
Db 2380 AGCGATGATGTACGGCTCTCTGACAGCAGCATGCCCATGCCCCATCTGCTCTGCCCTCTTG 2439  
QY 2623 TTTACAAGCCTCAAGTGTCTCAATGGTGTGAGAGCCAGGACACACAGCTTTGTGCTTTGTGCTAGC 2682  
Db 2440 TTTACAAGCCTCAAGTGTCTCAATGGTGTGAGAGCCAGGAGCCACATGCTCTCTC 2499  
QY 2683 TTTCAAGTCCATCTAGCCCATCAAGCCTTTCTGACGACAGCTTTTGACAACCTTTATCTGG 2742  
Db 2500 TTTCAAGTCCATCTAGCCCATCAAGCCTTTCTGACGACAGCTTTTGACAACCTTTATCTGG 2559  
QY 2743 GAGTTTTTCAGAACTGTCTTCAAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2802  
Db 2560 GAGTTTTTCAGAACTGTCTTCAAGTAGTTAGTTCAAGTGAACAGAGGGTCTTCCAGTTT 2619  
QY 2803 GGAGAAAAAGAGGTTTCCAGAGTGAATTTTAGCATAACTCAATTCGTAAAGAAATCTTGG 2862  
Db 2620 GGAGAAAAAGAGGTTTCCAGAGTGAATTTTAGCATAACTCAATTCGTAAAGAAATCTTGG 2679  
|||||

QY	2863	ACTTGAGCACCTAATGGATATATTTTGAGAGAGACAGATCACTTTGGATGTATTAGTTGA	2922
DB	2680	ACTTGAGCACCTAATGGATATATTTTGAGAGAGACAGATCACTTTGGATGTATTAGTTGA	2739
QY	2923	GATGGGGCACAGGAGCTGAAGAGATTGGAAATCAATGCTTTATGGACATAGGCACAACT	2982
DB	2740	GATGGGGCACAGGAGCTGAAGAGATTGGAAATCAATGCTTTATGGACATAGGCACAACT	2799
QY	2983	AAATTAAGGAGTCGACAGACTTATCTCCGGACACCAAGGCTTAACCCATATTTAACTTT	3042
DB	2800	AAATTAAGGAGTCGACAGACTTATCTCCGGACACCAAGGCTTAACCCATATTTAACTTT	2859
QY	3043	GAACACCTCTGGTAGTGGAAACAATTTTATAGATCTGTCTCTGATGATAAAGAGTTTCA	3102
DB	2860	GAACACCTCTGGTAGTGGAAACAATTTTATAGATCTGTCTCTGATGATAAAGAGTTTCA	2919
QY	3103	GTCTGGAGGAAGAGATCAAAGTACAGTTGCGAGAGCACAGAGATGGAGGTCATGCAGG	3162
DB	2920	GTCTGGAGGAAGAGATCAAAGTACAGTTGCGAGAGCACAGAGATGGAGGTCATGCAGG	2979
QY	3163	TGGNACTTCAACAGATACAATATTTCTCAAGATTCAGAAGGTTTGTACAGAANACTATG	3222
DB	2980	TGGNACTTCAACAGATACAATATTTCTCAAGATTCAGAAGGTTTGTACAGAANACTATG	3039
QY	3223	GGAAAGATACACTCACCGGAGAAAAGAAGTTCTTCAAGAAAACACCAACATGCGCAATGA	3282
DB	3040	GGAAAGATACACTCACCGGAGAAAAGAAGTTCTTCAAGAAAACACCAACATGCGCAATGA	3099
QY	3283	ACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATATTCACAAAAGGCTTTGTATGA	3342
DB	3100	ACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATATTCACAAAAGGCTTTGTATGA	3159
QY	3343	AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGGTGAANAACCTCTTC	3402
DB	3160	AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGGTGAANAACCTCTTC	3219
QY	3403	CAAAAGCAATCAATATGTATATGGAATTCGAGGAGGTACTGGGTGTCAGTTTCAACAAGA	3462
DB	3220	CAAAAGCAATCAATATGTATATGGAATTCGAGGAGGTACTGGGTGTCAGTTTCAACAAGA	3279
QY	3463	CAGATCTTGTGTACATTTGCCACAGGCAGCTGCTCTTTTCCGGGTGAACCTTTGGGAAAAGTC	3522
DB	3280	CAGATCTTGTGTACATTTGCCACAGGCAGCTGCTCTTTTCCGGGTGAACCTTTGGGAAAAGTC	3339
QY	3523	TTTCCTCGAGTTCAGTGCAATGAAATGSCAATTTCTCCTCCAGGTCACTCACTCAGTCCAC	3582
DB	3340	TTTCCTCGAGTTCAGTGCAATGAAATGSCAATTTCTCCTCCAGGTCACTCACTCAGTCCAC	3399
QY	3583	TGGTAGGCCCAAGTGAATAGGCCCTAGCATTTAGCTCAATATGTTATTACACAGAGAGAACA	3642
DB	3400	TGGTAGGCCCAAGTGAATAGGCCCTAGCATTTAGCTCAATATGTTATTACAGAGGAGAACA	3459
QY	3643	GGCTTATCCTGAGTATTTAAATTAATTAATTAATTAATGAGGCTGAAGGTATGGTGCATGG	3702
DB	3460	GGCTTATCCTGAGTATTTAAATTAATTAATTAATTAATGAGGCTGAAGGTATGGTGCATGG	3519
QY	3703	ATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAAAGCAGCAGTGG	3762
DB	3520	ATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAAAGCAGCAGTGG	3579
QY	3763	CCTCTACGTTTTTACTCTCTTTGCTGAAAAAAA	3794
DB	3580	CCTCTACGTTTTTACTCTCTTTGCTGAAAAAAA	3611

## RESULT 2

US-09-350-982C-4

; Sequence 4, Application US/09350982C

; Patent No. 6455290

; GENERAL INFORMATION:

; APPLICANT: Berthelsen, Jens

APPLICANT: Toma, Salvatore

QY 1063 TAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAGTTATGGTGCAGACCCAACTGCT 1122  
DB 858 TAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAGTTATGGTGCAGACCCAACTGCT 917  
QY 1123 CAATTTGCACAAATAAAGTCTATAGACTTGGCTCCACACACAGTAAAGAAAGATT 1182  
DB 918 CAATTTGCACAAATAAAGTCTATAGACTTGGCTCCACACACAGTAAAGAAAGATT 977  
QY 1183 AGCATATGAATTTAAAGGCCACATCGTTGCTGCAAGTGCAGGAGAGCTGATGTTACTCG 1242  
DB 978 AGCATATGAATTTAAAGGCCACATCGTTGCTGCAAGTGCAGGAGAGCTGATGTTACTCG 1037  
QY 1243 AATCAAAAAACATCTCTCTCTGGAATGTTGAATTTCAAGCATCTCCAAACACATGAAC 1302  
DB 1038 AATCAAAAAACATCTCTCTCTGGAATGTTGAATTTCAAGCATCTCCAAACACATGAAC 1097  
QY 1303 AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAGAGCTAAAGTATGTTGAATGTT 1362  
DB 1098 AGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAGAGCTAAAGTATGTTGAATGTT 1157  
QY 1363 GCTAAGAAAAAGGAGCAACATCAATGAAAGACTAAAGAAATCTTGACTCTCTGCAAGT 1422  
DB 1158 GCTAAGAAAAAGGAGCAACATCAATGAAAGACTAAAGAAATCTTGACTCTCTGCAAGT 1217  
QY 1423 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT 1482  
DB 1218 GGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAAAGT 1277  
QY 1483 TAATGCTCTGGAATAATCTGGTGCAGACTCTCTACAGAGCTGCATATGTTGGTGCATCT 1542  
DB 1278 TAATGCTCTGGAATAATCTGGTGCAGACTCTCTACAGAGCTGCATATGTTGGTGCATCT 1337  
QY 1543 ACAAGCTGCGGCTACTCTGAGCTATGGGTGTGATCTTAACATTAATATCCCTTCAGGG 1602  
DB 1338 ACAAGCTGCGGCTACTCTGAGCTATGGGTGTGATCTTAACATTAATATCCCTTCAGGG 1397  
QY 1603 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCACTCTCCAGAGAGGTATCTC 1662  
DB 1398 CTTTACTGCTTTACAGATGGAAATGAAATGTACAGCACTCTCCAGAGAGGTATCTC 1457  
QY 1663 ATTAGTAAATTCAGAGGCACAGACAATTTGCTGGAAGTGCAGAGCTGGAGATCTCGA 1722  
DB 1458 ATTAGTAAATTCAGAGGCACAGACAATTTGCTGGAAGTGCAGAGCTGGAGATCTCGA 1517  
QY 1723 AACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCACTGCAGAGACATTTGAAGGGCTCA 1782  
DB 1518 AACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCACTGCAGAGACATTTGAAGGGCTCA 1577  
QY 1783 GTCTACACACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1842  
DB 1578 GTCTACACACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCT 1637  
QY 1843 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCTTGTACCTTTGACACAATGC 1902  
DB 1638 ACAGCATGGAGCTGATGTCATGCTAAAGATAAAGRSGCTTGTACCTTTGCACAATGC 1697  
QY 1903 ATGTTCTTTAGACATATTAAGTGTCCAGACTCTCTGTTTAAACATGAGAGCTAGTTAA 1962  
DB 1698 ATGTTCTTTAGACATATTAAGTGTCCAGACTCTCTGTTTAAACATGAGAGCTAGTTAA 1757  
QY 1963 TGTAGCTGATTTATGAAATTTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATGA 2022  
DB 1758 TGTAGCTGATTTATGAAATTTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATGA 1817  
QY 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAAA 2082  
DB 1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAAA 1877  
QY 2083 TACTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGA 2142  
DB 1878 TACTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGA 1937

QY 2143 TGCAGCTTTTCTAGATGCTGCCAAGAGGGTTGTTTACCCAGAGTGAAGAGTTGTTCTTC 2202  
DB 1938 TGCAGCTTTTCTAGATGCTGCCAAGAGGGTTGTTTACCCAGAGTGAAGAGTTGTTCTTC 1997  
QY 2203 TCTGTATAATGTTAAATGTCGCGGATACCCAGGAGACATTTCAACACCTTTTACATTTAGC 2262  
DB 1998 TCTGTATAATGTTAAATGTCGCGGATACCCAGGAGACATTTCAACACCTTTTACATTTAGC 2057  
QY 2263 AGCTGGTTATAAATAATTTAGAAAGTTGCAGAGTATTTGTTTACACACAGGAGTGAAGTGA 2322  
DB 2058 AGCTGGTTATAAATAATTTAGAAAGTTGCAGAGTATTTGTTTACACACAGGAGTGAAGTGA 2117  
QY 2323 TGCCCAAGACAAAGAGGAGGACTTTTCTTTTACATAATGTCAGCATCTTACGGGATGAGA 2382  
DB 2118 TGCCCAAGACAAAGAGGAGGACTTTTCTTTTACATAATGTCAGCATCTTACGGGATGAGA 2177  
QY 2383 TGTAGACGCTCTACTTAATAAAGTATTAATGATGTTGTCATATGCCACGACAAATAGGGCTTT 2442  
DB 2178 TGTAGACGCTCTACTTAATAAAGTATTAATGATGTTGTCATATGCCACGACAAATAGGGCTTT 2237  
QY 2443 CACACCTTTTGCAGAGCAGCCCAAGGAGGACGACACACAGCTTTGCTTTGTTGCTAGC 2502  
DB 2238 CACACCTTTTGCAGAGCAGCCCAAGGAGGACGACACACAGCTTTGCTTTGTTGCTAGC 2297  
QY 2503 CCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTTAGATTTAGTTTC 2562  
DB 2298 CCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTTAGATTTAGTTTC 2357  
QY 2563 AGCGGATGATGTCAGGCTCTTCTGACAGAGCCTATGCCCCCATCTGCTCTGCCCTCTTG 2622  
DB 2358 AGCGGATGATGTCAGGCTCTTCTGACAGAGCCTATGCCCCCATCTGCTCTGCCCTCTTG 2417  
QY 2623 TTACAAGCTCAAGTGTCTCAATGTTGTGAGAAGCCAGGAGGACCTGCACATCTCTCTC 2682  
DB 2418 TTACAAGCTCAAGTGTCTCAATGTTGTGAGAAGCCAGGAGGACCTGCACATCTCTCTC 2477  
QY 2683 TTCAGTTCCTATCTAGCCCTTCTGACAGCAGCTTTCTGACGAGCAGCTTTGACAACTTATCTGG 2742  
DB 2478 TTCAGTTCCTATCTAGCCCTTCTGACAGCAGCTTTCTGACGAGCAGCTTTGACAACTTATCTGG 2537  
QY 2743 GAGTTTTCAGACTCTCTCTCAGTAGTTAGTTTCAAGTGGAGAGAGGGTCTTCCAGTTT 2802  
DB 2538 GAGTTTTCAGAACTCTCTCTCAGTAGTTAGTTTCAAGTGGAGAGAGGGTCTTCCAGTTT 2597  
QY 2803 GGAGAAAAAGGAGGTTTCCAGGAGTAGATTTTAGCATAAATCTCAATTCGTAAAGAACTCTGG 2862  
DB 2598 GGAGAAAAAGGAGGTTTCCAGGAGTAGATTTTAGCATAAATCTCAATTCGTAAAGAACTCTGG 2657  
QY 2863 ACTTGAGCCTTAATGATATATTTGAGAGAGAACAGATCAGCTTTGGATGATTTAGTTGA 2922  
DB 2658 ACTTGAGCCTTAATGATATATTTGAGAGAGAACAGATCAGCTTTGGATGATTTAGTTGA 2717  
QY 2923 GATGGGCAACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2982  
DB 2718 GATGGGCAACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACT 2777  
QY 2983 AATTAAGAGGTCGAGAGCTTATCTCCGACAAAGGCTTTAAACCCATATTTAACTTT 3042  
DB 2778 AATTAAGAGGTCGAGAGCTTATCTCCGACAAAGGCTTTAAACCCATATTTAACTTT 2837  
QY 3043 GAACACTCTGTTAGTGGACAAATTTCTTATAGATCTCTCTCTGATGATTAAGAGTTTCA 3102  
DB 2838 GAACACTCTGTTAGTGGACAAATTTCTTATAGATCTCTCTCTGATGATTAAGAGTTTCA 2897  
QY 3103 GTCTGTGGAGGAGAGATGCAAGTACAGTTCGAGAGCAGAGATGGAGGTCATCGAGG 3162  
DB 2898 GTCTGTGGAGGAGAGATGCAAGTACAGTTCGAGAGCAGAGATGGAGGTCATCGAGG 2957  
QY 3163 TGAATCTTTCAACAGATACAATATTTCTCAAGATTCAGAGGTTTGTAAACAAGAACTATG 3222  
DB 2958 TGAATCTTTCAACAGATACAATATTTCTCAAGATTCAGAGGTTTGTAAACAAGAACTATG 3017  
QY 3223 GGAAAGATACACTCAGGAGAGAAAGAAAGTTTCTGAGAGAAACCAACCACTCCCAATGA 3282



QY 1283 CATCCTCAAAACACATGAACAGCATTGCAFTTGTGCTGCATCTCCATATCCAAAGA 1342  
D 1557 CACCCGAGTCTCATGAACAGACACTGCACCTGTGCTGGCCCTCTGCACTCCCAACGT 1616  
QY 1343 AAGCAAAATATGTGAAGTGTGCTTAAGAAAGAGGACAAACATCAATGAAAGACTAAAGAA 1402  
D 1617 AAACAAGTGACAGAAATGTTACTTAGAAAGAGGACAAATGTTAATGAAAAAAATAAGAT 1676  
QY 1403 TTCCTTCACTCCTCTGCACGTGGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTG 1462  
D 1677 TTCATGACTCCCTCGATGTGTCAGCGGAAAGAGCCCATTAATGATGTCATGGAAGTTCTG 1736  
QY 1463 GTGAACATCAAGCAAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522  
D 1737 CATAAGCATGGCCCAAGATGAATGACCTGACACCCCTTGGTCAGACTGCTTTGCAATAGA 1796  
QY 1523 GCTGCATATTTGGTGCATCTACAAACCTGCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582  
D 1797 GCGGCCCTAGCAGGCCACCTGCAGACCTGCGCCTCTGCTGAGTTACGGCTCTGACCCC 1856  
QY 1583 ACATTTATATCCTTCCAGGCTTACTGCTTACAGATGGAAATGAAATGTACAGCAA 1642  
D 1857 TCATCATCTCTTCAAGGCTTCAGAGCAGACAGATGGCAATGAAGCAGTGCAGCAG 1916  
QY 1643 CTCCTCCAAGAGGGTATCTCAATAGTAGTAATTCAGAGGCAGACAGCAATTTGCTGGAAGCT 1702  
D 1917 ATTCTGAGTGAGAGTACACCTATACCTACTTCIGATGTTGATTCGACTCTTAGAGGCA 1976  
QY 1703 GCMAAGCTGGAGATGCGAACTGTAAAAAACCTGTGACTGTTCAGAGTGTCAACTGC 1762  
D 1977 TCTAAAGCTGGAGACTGTGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATGT 2036  
QY 1763 AGAGCATTTGAAGGGCTGAGTCACTACACACTTCATTTTGCAGCTGGGTATACACAGTG 1822  
D 2037 AGAGACTTAGAGGGCGGCTTCCAGCCCTTACACTTCGCAGCAGGCTACAAACCGCTG 2096  
QY 1823 TCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCG 1882  
D 2097 TCTGTGTAGAGTACTGCTACACACGCTGCCGATGCTCCATGCCAAAGCAAGGCTGC 2156  
QY 1883 CTTGTACCTTTCCAAATCATGTTCTTATGACATATGAAAGTGTGCAAGACTTCTTGT 1942  
D 2157 TTGGTGGCCCTTCATAAGCCCTGTCATAGACACTATGAGTGGCTGAGCTTTTAGTA 2216  
QY 1943 AAACATGGAGCAGTAGTAAATAGTCTGATTTATGAAATTTACACCTTTACATGAAGCA 2002  
D 2217 AGGCATGGGCTTCTGCAATGTGGCGACTTATGGAATTTACCCCTCTCCATGAAGCA 2276  
QY 2003 GCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACC 2062  
D 2277 GCAGCTAAAGGAAATATGAAATCTGCAAGCTCTCTTTTAAACATGGAGCAGATCCAAC 2336  
QY 2063 AAAAAACACAGGATGGAATACTCTTTGGATCTGTTAAAGATGGAGATACAGATATT 2122  
D 2337 AAAAGACAGAGATGGAATACTCTTTGGATTTGGTAAAGAAAGAGACACAGATATT 2396  
QY 2123 CAAGATCTGCTTAGGGGAGATGACCTTGTCTGATGCTGCCAAGAGGGTTGTTTAGCC 2182  
D 2397 CAGGACTTACTGAAAGGGATGCTCTCTTTGTTGGATGCTGCCAAGAGGGCTGCTGGCA 2456  
QY 2183 AGAGTGAAGAGTCTCTCTGCTGATTAATGTAATTTGCCCGGATPACCAAGGCAGACAT 2242  
D 2457 AGAGTGAAGAGTCTGTAACCCCAAGAGATATCAACTGCAGAGACACCCAGGCGAGAAAT 2516  
QY 2243 TCAACACCTTTACATTTAGCAGCTGTTTATATATTTAGAGTTGCAAGATATTGTTA 2302  
D 2517 TCAACCCCTCTGCACCTGGCAGCGGCTATATAACCTTGGAGTAGCTCAATATCTCTA 2576  
QY 2303 CAACAGGAGCTGATGTGAATGCCCAAGACAAAGAGGAGCTTATTCCTTTACATAATGCA 2362  
D 2577 GAGCATGGAGCTGATGTTAATGGCCAGCACAAGGGTGTGTTAAATTCCTCTTCATANTCGG 2636

QY 2363 GCATCTACGGGCATGTAGATGTAGCAGCTCTACTATAAAGATATATGATGTGTCAAT 2422  
D 2637 GCATCTTATGGGCATGTGACATAGCGGCTTTATTTGATAAAATACAAACAGTGTGTAAAT 2696  
QY 2423 GCACAGCAAAATTTGGCTTTTCACACCTTTTCACGAAAGCAGCCCAAAAGGAGGCAACACAG 2482  
D 2697 GCAACAGATAAGTGTGGGCTTTACTCCCTCCATGAAGCAGCCCAAGAAAGGAGGCGCAG 2756  
QY 2483 CTTTGTGCTTTGTTGTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAA 2542  
D 2757 CTGTGCGCCTCTCTCTAGCGCATGTGCAACCCACCCACCATGAAGAACCCAGGAAGGCCAG 2816  
QY 2543 ACACCTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCC 2602  
D 2817 ACOCCTCTGATCTGCAACAGACTGACATATCAGAGCTTTGCTGATAGATGCTCATGCC 2876  
QY 2603 CCATCTGCTCTGCCCCCTTTTACAAGCCCTCAAGTGTCTAATGTTGTGAGAAGCCAGGA 2662  
D 2877 CCAGAGGCCCTTACCTACCTGTTTAAACCTCAGG-----CTACT 2915  
QY 2663 GCCACTGCAGATGCTCTCTCTCAGTGCATCTIAGCCCATCAAGCCCTTTCTGACGCGCAG 2722  
D 2916 GTAGTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCTCTGCTCGGCTGCCAGC 2975  
QY 2723 AGTCTTGAACAACCTATCTGGAGTCTTTTTCAGAACTCTCTTCACTAGTTAGTTCAAGTGA 2782  
D 2976 AGCATAGACAACCTCTACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCCTCCAATGCA 3035  
QY 2783 ACAGAGGGTCTTCCAGTTTGGAGAAAAAG-----GAGTTCACAGGAGTAGATTTTAGC 2836  
D 3036 GGGATGGCGCGCGGGAACAGAAAGGAAGAGAGAGAGTTGCTGCTTGTGACATGAAT 3095  
QY 2837 ATAACATCAATTCGTAAGGAATCTTGACCTTGAGCAGCTTAATGGATATATTTGAGAGAGAA 2896  
D 3096 ATCAGCCAAATTTCTAAAAAGCCTTGGCCTTGAACACTCTCGGATATCTTTGAAACAGAA 3155  
QY 2897 CAGATCACCTTTGGATGATTTAGTTGAGATGGGCAAGAGAGCTGAAGAGATTTGGAATC 2956  
D 3156 CAGATTACACTAGATGTTGGCTGATATGGTTCATGAAGAGTTGAAGAAATATGAGCATC 3215  
QY 2957 AATGCTTATGAGCATAGGCACAAACTTAATAAGGAGTGCAGAGACTTATCTCCGAGCAA 3016  
D 3216 AATGATATGGCAGCGCGCACAAATTAATCAAGGAGTAGAAGACTCTTAGTGGACAA 3275  
QY 3017 CAAGTCTTAAACCCATATTTAACTTTGAACCTCTGGTAGTGGAGCAATTTCTTATAGAT 3076  
D 3276 CAAGGACCAATCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGTGGAT 3335  
QY 3077 CTGCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCA 3136  
D 3336 CTTGCTCCAGAGATATAAGAAATATCAGTCAGTGGAGAGAGATGCAAAAGTACTATTCA 3395  
QY 3137 GAGCAGAGATGGAGGTCATCAGTGGAAATCTTCAACAGATACAATATTTCTCAAGATT 3196  
D 3396 GAAACAGAGATGGTGGTAAATGCTGGCGCATCTTCAACAGATACAATGTCTATCGAATT 3455  
QY 3197 CAGAAGGTTTGAACAAGAACTATGGGAAACATACACTCACCGGAGAGAAAGAGTTCT 3256  
D 3456 CAAAAGTTGTCAACAGAAAGTTGAGGAGCGGTTCTGCCACCCAGAGAAAGGTCT 3515  
QY 3257 GAAGAAAACCAACCATGCCCATAAGCAATGCTATTTCATGGGTCTCTTTTGTGAAAT 3316  
D 3516 GAGGAGATCAACACCATCACAATCAGCGCATGTTGTTTCATGGTTCCTTTTCATTAAT 3575  
QY 3317 GCAATTATCCAAAGGCTTTGATGAAAGCATGGTACATAGGTGGTATGTTTGGAGCT 3376  
D 3576 GCAATTATTCATAAGGCTTTGATGAGGACATGATACATAGGAGGAATGTTTGGGGCC 3635  
QY 3377 GGCATTTATTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGATATGGAATTCGAGA 3436  
D 3636 GGGATTATTTGCTGAAAACCTCTTCCAAAAGCAACCAATATGTTTATGGAATTCGAGA 3695  
QY 3437 GGTACTGGGTGTCCAGTTTCACAAAGACAGATCTTGTATACATTTGCCACAGGCGTCTC 3496

Db 3696 GGAACAGGCTGCCCTACACAAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTC 3755  
QY 3497 TTTTCCCGGGTAACCTTCGGGAAAGTCTTTCTCGAGTTCAGTGAATCAAAATGGCACAT 3556  
Db 3756 TTCTGTAGAGTGACCTTCGGGAAATCCCTTCGAGATTACACATGAAATGGCCAC 3815  
QY 3557 TCTCTCCAGGTCATCACTCACTGCTAGTGGCCAGTGTAAATGGCTTAGCATTTAGCT 3616  
Db 3816 GCGCTCCAGGSCACCACTCAGTCATTGTTAGACCGAGCGTCAATGGCTGGCATATGCT 3875  
QY 3617 GAATATGTTTACAGAGGAGACAGCTTATCTGAGTATTTAAATTTACTTACCAAGATT 3676  
Db 3876 GAATATGTCATCTACAGAGGAGACAGCATACCCAGAGTAGTCTTATCACTTACCAGATC 3935  
QY 3677 ATGAGGCTGGAAG 3689  
Db 3936 ATGAAGCCAGAG 3948

RESULT 4  
US-09-196-387-7  
; Sequence 7, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196.387  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/095.225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-3800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 6..2027  
US-09-196-387-7

Query Match 38.2%; Score 1449.6; DB 4; Length 4491;  
Best Local Similarity 65.3%; Pred. No. 0;  
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;

QY 383 TCCTCTGGCTTTTATCATTAAGGTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442  
Db 657 TCTCCCTGCACCTCGCTGCAGCTTTTGGAAAGCAAGGATGTTCTAGAACACTTTACTACAG 716  
QY 443 AATGGTGAAGTGTCAAGACACGTGATGATGGGGGCTTATTCCTCTTCTATATGATGC 502  
Db 717 ATGGGTGCTAATGTCACGCTCGTATGATGAGGCTCTCATCCCGTTCTCATATATGCTGT 776  
QY 503 TCTTTTGGTGCATGCTGAAGTAGTCAATCTCTTTTTCGACATGTCGACACCCCAATGCT 562  
Db 777 TCTTTTGGCCATGCTGAGGTTGTAGTCTGTTATTGTGCAAGAGAGCTGATCCAAATGCC 836  
QY 563 CGAGATAAATTTGGAATTTATCTCTCTCCATGAAGCTGCAATTTAAAGGAAGATTGATGTT 622  
Db 837 AGGATAAATGGAACATATACACCTCTGTCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896  
QY 623 TGCATTTGCTGTTTACAGCATGGAGCTGAGCCCAACCATCGGAATATACAGATGGAAGGACA 882  
Db 897 TGCATTTGCTGCTGTCAGCAGCGAGGCTGACCCAAACATTCGGAACACTGTATGGGAATCA 956  
QY 683 GCATTTGATTTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGTGAATATAAGAAAGAT 742  
Db 957 GCGCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCTTACAGTGAATACACAAAGAC 1016  
QY 743 GAATCTTTAGAAAGTGCAGAGTGGCAATGAAGAAAAATGATGGCTCTTACTCACACCA 802  
Db 1017 GAATCTCTAGAAGCTGCTAGSAGTGGTAAATGAAGAAAACTAATGCTTTTACTGACTCCT 1076  
QY 803 TTAATCTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCCATTACATTATTTGGCAGCA 862  
Db 1077 CTAATGTGAATTTGCCATGCAAGTGTGGCGAAAGTCGACTCTCTTTACATCTAGCAGCG 1136  
QY 863 GGATATAACAGAGTAAAGATTGTACAGCTCTTACTGCAACATGGAGCTGATGCTCCATGCT 922  
Db 1137 GGCTACACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGCTGCTGAIGTTCAFGCA 1196  
QY 923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTTATGGTCAATTTAAGTA 982  
Db 1197 AAAGACAAGGTGGACTTTGCGCTTTCATAATGCATGTTTCATATGGACATTTATGAAGTC 1256  
QY 983 ACTGAATTTTGGTCAAGCATGGTGTGTAATGAATGAAGTGGAGTGTGGCAATTCACAT 1042  
Db 1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316  
QY 1043 CCTCTTCATGAGCAGCTTCTAAGAACAGGTTCAAGATATGTTCTCTCTCTTAAGTTAT 1102  
Db 1317 CCACTGCACGAGGCTGCTTCCAGAACCGTGTAGAGTCTGCTCTTTGTTACTTAGCCAT 1376  
QY 1103 GGTGCAGACCCCAACACTGCTCAATTTGTCACAATAAAAAGTGTATAGACTTTGGCTCCCA 1162  
Db 1377 GCGCTCATCTCCTACGTTAGTCACTGCCATGGCAAAAGTGTGTGGATATGGCTCCAACT 1436  
QY 1163 CCACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222  
Db 1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTTCTTTACTACAGCAGCC 1496  
QY 1223 CGAGAAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGAAATGGTGAATTTTCAAG 1282  
Db 1497 AGAAGCAGACTTAGCTAAAGTTAAAAAACACACTCGCTCTGGAAATCATTAATTTCAA 1556  
QY 1283 CATCTCAAAACACATGAACAGCATTTGCTGCTGCTGCTATCCATATCCCAAAAGA 1342  
Db 1557 CAACCGCAGTCTCATGAACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1616  
QY 1343 AAGCAAAATATGTAACCTGTTGCTTAAGAAAGGAGGACCAACATCAATGAAAGACATAAGAA 1402  
Db 1617 AACAAAGTGACAGAAATTTGTTACTTTAGAAAAGGAGCAATGTTTATGAAAAAATAAGAT 1676  
QY 1403 TTCTTTGACTCCTCTGACGCTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTG 1462  
Db 1677 TTTATGACTCCCTGTCATGTTGTCAGCGGAAAGGCCCATATGATGTCATGGAAGTTCTG 1736





Db 3876 GAGAAACACACATCACAATGAGCGCATGTGTTTCATGGTCTCTCTTCATTAATGCC 3935  
Qy 3320 ATTATCCACAAGGCTTTGATGAAGCATCCGTACATAGTGTGATGTTGGAGCTGGC 3379  
Db 3936 ATTATTCATAAAGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGCGGG 3995  
Qy 3380 ATTTATTTTGTGAAAACTCTTCCAAAAGCAATCAATATATATATGGAATTTGGAGGAGT 3439  
Db 3996 ATTTATTTTGTGAAAACTCTTCCAAAAGCAATCAATATATGTTATGGAATTTGGAGGAG 4055  
Qy 3440 ACTGGGTGTCAGTTTCAAAAGCAGATCTTTGTTACATTTGTCACAGGAGCTGCTCTTT 3499  
Db 4056 ACAGGCTGCCCTACACACAGGAGGATGCTCTATATATGTCACAGCAAAATGCTCTTC 4115  
Qy 3500 TGCCGGGTAACTTTGGAAAGCTTTCTCGAGTTCAGTCAATGAAATGGACATTTCT 3559  
Db 4116 TGTAGTGCACCTTGGAAATCTTTCTCGAGTTTACACCATGAAATGGCCAGCGG 4175  
Qy 3560 CTTCCAGGTATCACTCAGTCACTGGTAGGCCAGTGTAAATGGCCTAGCATTAAGCTGAA 3619  
Db 4176 CTTCCAGGCACTCAGTCACTGGTAGGCCAGTGTAAATGGCCTAGCATTAAGCTGAA 4235  
Qy 3620 TATGTTATTTACAGAGAGACAGCTTATCTCTCAGTATTTAATTTACTTACCAGATTATG 3679  
Db 4236 TATGTCATCTACAGAGGAGACAGGATACCCAGAGTATCTTATCACTTACCAGATCATG 4295  
Qy 3680 AGGCCTGAAG 3689  
Db 4296 AAGCCAGAG 4305

RESULT 5

US-09-196-387-9  
; Sequence 9, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/095,225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4657 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double

; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 6..2855  
US-09-196-387-9  
Query Match 34.2%; Score 1297.8; DB 4; Length 4657;  
Best Local Similarity 62.9%; Pred No. 0;  
Matches 2400; Conservative 0; Mismatches 907; Indels 508; Gaps 5;  
Qy 383 TCCTCTGCTGCTTATCATTAAGGTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAG 442  
Db TCCTCCCTGCACTTCGCTGCAAGGTTTGGAAAGGAGGATGTTGTAGAACACTTACTACAG 716  
Qy 443 AATGGTGCAGAGTGTCAAGCAGCTGATGGGGCTTTATTCCTCTTTCATATGCATGC 502  
Db ATGGGTGCTAATGTCACGCTGCTGATGATGGAGTCTCATCCCGCTTCATATGCCTGT 776  
Qy 503 TCTTTTGGTTCATGCTGAAGTAGTCAATCTCTTTTTCGACATGGTGCAGACCCCAATGCT 562  
Db TCTTTTGGCCATGCTGAGGTTGTGAGTCTGCTTATTTGCGCAAGGAGCTGATCCAAATGCC 836  
Qy 563 CGAGATAATTTGGAATTTATATCTCTCTCCATGAAGCTGCAATTAAGAGAAAGATTTGATGT 622  
Db AGGATAAATCTGGAATTTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896  
Qy 623 TGCAATTGCTGTTTACAGCATGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACA 582  
Db TGCAATTGCTGCTGTCGACGACGAGCTGACCCCAACCATTCGGAACACTGATGGGAAATCA 956  
Qy 683 GCATTGGAATTTAGCAGATCCATCTGCCAAAGCAGTCTTACTGCTGAATATAAGAAAGAT 742  
Db GCCTTGACCTGCGAGATCTCTTCAGCAAAAGCTGCTCTACAGCTGAATACAGAAAGAC 1016  
Qy 743 GAACTCTTAGAAAAGTCCAGGAGTGGCAATGAAGAAAATGATGCTCTCTACACACCA 802  
Db GAACTCTTAGAAAAGTCCAGGAGTGGTGAATGAAGAAAATGATGCTCTCTACACACCA 1076  
Qy 803 TTAATCTCAACTGCCACGCAAGTGGCAGAAAGTCAACTCAATACATTTGGCAGCA 862  
Db TTAATCTCAACTGCCACGCAAGTGGCAGAAAGTCAACTCAATACATTTGGCAGCA 1136  
Qy 863 GGATATAACAGAGTAAAGATTTGACAGCTGTTACTGCAACATCGAGCTGATGCTCCATGCT 922  
Db GGATATAACAGAGTAAAGATTTGACAGCTGTTACTGCAACATCGAGCTGATGCTCCATGCT 1196  
Qy 923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCAATTAAGTA 982  
Db AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCAATTAAGTA 1256  
Qy 983 ACTGAACCTTTTGGTCAAGCATGGTGTGTAATGAATGCAATGGAGCTTGTGGCAATTCAC 1042  
Db ACTGAACCTTTTGGTCAAGCATGGTGTGTAATGAATGCAATGGAGCTTGTGGCAATTCAC 1316  
Qy 1043 CCTCTTCATGAGGAGCTTCTTAAGACAGAGGTTGAAGTATGTTCTCTCTCTTCTTAAAGT 1102  
Db CCTCTTCATGAGGAGCTTCTTAAGACAGAGGTTGAAGTATGTTCTCTCTCTTCTTAAAGT 1376  
Qy 1103 GGTGCAAGACCCCAACACTGCTCAATTTGACACAATAAAGTCTCTATAGACTTGGCTCCACA 1162  
Db GGTGCAAGACCCCAACACTGCTCAATTTGACACAATAAAGTCTCTATAGACTTGGCTCCACA 1436  
Qy 1163 CCACAGTTAAAGAAAGATTTAGCATATGAATTTAAAGGCCCACTGCTTGTGCAAGCTGCA 1222  
Db CCACAGTTAAAGAAAGATTTAGCATATGAATTTAAAGGCCCACTGCTTGTGCAAGCTGCA 1496  
Qy 1223 CGAGAAGCTGATGTTACTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAAT 1282  
Db CGAGAAGCTGATGTTACTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAAT 1556  
Qy 1283 CATCCTCAACACATGAAGAACAGCATTTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1342



[illegible]





Qy	1373	GGACCAACATCAATGAAAGACTAAGAATAATTCTTTGAOCTCCTGCAGTGGCATCTGAG	1438
Db	178	GGCACCCATATCAACACATGCATCGAATGCCATTATGCAATTCATGGCTGCCAAA	237
Qy	1433	AAAGCTCATATGATGTCTTGGAAGTAGTGGTGAACATGAAGCAAGGTTTAATGCTCTG	1492
Db	238	GAAAGTTCATCATGAAGTGGFCGGAACCTTCGAAAAAGAACAGATGTTGATGCTGCC	297
Qy	1493	GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGTCATCTACAACCTGC	1552
Db	298	ACTAGAAGGTAACACAGCGTTTACATATAGCATCATTTGGCAGNACAGAACTAATCGTC	357
Qy	1553	COCCTACTCCTGAGCTATAGGGTGTGATCCTAACTATATATCCTTCAGGGCTTTACTGCT	1612
Db	358	ACAGTACTTCTTGAATAAGTGGTAACTTAACTGTTAACTGACAACTAAACGGTTTTACACCA	417
Qy	1613	TTACAGATGGAAATGAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGTAAAT	1672
Db	418	CTTTTACATGGCTGCACAGAAATCAGAAATCTGTTGTACGCTATCTTCTTGCCACCACAAT	477
Qy	1673	TCAGAGCGACAGACAATTTGCTGAAGCTGCAAAAGGCTGGAGATGT-----CGAATAC	1725
Db	478	GCCAATCAAGCTTTAAGTACAGAGACGGTTTACGCCACTGGCAGTTTGCCTTCCACACAA	537
Qy	1726	TGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAAGTC	1785
Db	538	GGTCACGATGCTGTGCTGCTGTTTTGCTTTGAAATGACACGCGCGGGAAGTG-CGCTT	596
Qy	1786	TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGSANATATCTGCTACA	1845
Db	597	GCCAGCACTGCATTTGCTGTCTAAAAAGATGATACGAAAGCAGCTACGCTATATTACTTCA	656
Qy	1846	GCATGGAGCTGATGTGCATGCTAAAGATAAGAGAGCCCTTGCTAGCTTGCACAAATGCATG	1905
Db	657	AAATGAGCATAACTCGATGTGCATTCGNAAGCGCTTTTACTCCGCTTCATATCGCCGC	716
Qy	1906	TTCTTATGGACATTTAAGTTGCGAACTTCTTTGTTAAACATGGACAGTAGTTAATGT	1965
Db	717	TCACTATGAAATGACAACGTAGCAACTGCTACTCGAAAAGGAGCCAAATGTGAATTA	776
Qy	1966	AGCTGATTTATGAAATTTACACCTTTACATCAAGCAGCACCAAAGGAAAAATATGAAAT	2025
Db	777	CCAAAGCAGACATACATAAAGTCGTTACAGTTTGCACAAAAATGGGGTCGTACAAACAT	836
Qy	2026	TTGCAAACTTCTGCTCCAGCATGGTCGAGACCCCTACCAAAAAAACAGGATGGAATAC	2085
Db	837	GGTTTCGTTATCTTGGCTCATGGGCCGTAATTCAGCTGTCGCACAGTATTTACTAAC	896
Qy	2086	TCCTTTGGATCTTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGTTAGGGAGA	2142
Db	897	ACCAATTACACTGTGCTCTGCTTCAGGTCATGATCAAGTCTTTCATTTGTTGCTTGAANA	956
Qy	2143	TGCAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTTACCCAGAGTGAAGAGTTGCTTC	2202
Db	957	AGGAGCTCCAATCAGTGCTTAAGCAAAAAATTTGGTTGGCTCCCTTACATATGCGACGACA	1016
Qy	2203	TCCTGTATGTAAATTTGCCGGCATACCCAAGGCAGACATTCACACACCTTTACATTTAGC	2262
Db	1017	GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCAITGGC	1061
Qy	2263	AGCTGTTTATAAATTTAGAAGTTGCAGAGTATTTGTTACCAACAGCGACTGATGTAA	2322
Db	1062	TGCTCATTCGGACATGTCCGTCGCTAAACTTTTGTGATGATGTAATCTGCACCCGAA	1121
Qy	2323	TGCCCAAGACAAGAGGACTTTATTCCTTTACATAANTGCAGCATCTTTACGGGATGTAGA	2382
Db	1122	TGCTCGAGCTTCAATGGCTTCCACACCGCTGCATCGCTTGCAAAAAANAATCGCATTA	1181
Qy	2383	TGTAGCAGCTCTACTAATAAAGTATAANTGCATGTGCAATGCCACGACAAATGGCTTT	2442
Db	1182	AATTGTGCACTGCTACTGAATACACGCTGCAATCGAAGCACTACTGAACTGGCTCT	1241
Qy	2443	CACACCTTTGCCAGAGCAGCCCCAAAAGGGACGAACAGCTTTGTCTTTGTGTGCTAGC	2502

```

Db 1242 CTCACCGCTGCATGCGCTGCTTTTATGGGTGCTATAACAATTGTCACTATTACTATACA 1301
QY 2503 CCATGGAGTGACCGGACCTCTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTAGTTTC 2562
Db 1302 ACAAGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAACGCCCTCTTCATTTAGCTGC 1361
QY 2563 A 2563
Db 1362 A 1362

RESULT 10
US-08-847-429A-36/c
; Sequence 36, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scott
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-847-429A-36

```

[illegible]





Db 538 GGTCCAGATCGTGGTGGTCTTTTCTTGAATAACACGCGCGGAAAGTG-CGCTT 596  
QY 1786 TACACCATTCTTTTCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGTACA 1845  
Db 597 GCCAGCACTGCATATCTCTAAAGAGATGATACAAAGCAGCTACGCTATTACTTCA 656  
QY 1846 GCATGGAGCTGATGTCATGCTAAAGATAAAGGAGGCTTGTACCTTTGGCACAATGATG 1905  
Db 657 AAATGAGCATAAATCGCATGTGACTTCCGAAAGCGGCTTACTCCGCTTCATATCGCCGC 716  
QY 1906 TTCTTATGACATATCAAGTTCAGACACTTCTTGTAAACATGAGCAGCAGTAGTAATGT 1965  
Db 717 TCACATGGAATGAGAACGTGACACACTGCTACTCGAAAGGAGCCCAATGTGAATTA 776  
QY 1966 AGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAAGCAAAATATGAAT 2025  
Db 777 CCAAGCAGACATAACATAAGTCGGTTACAGTTGCAACAAATGGGTGCTACAAACAT 836  
QY 2026 TTGCAAACTTCTGTCAGCATGGTGCAGACCTTACCAAAAAAAGGAGGATGGAATAC 2085  
Db 837 GGTTCGTTATGTTGGCTCATGGGCGGTAAATGACTGTCGCACACCTGATTTACTAAC 896  
QY 2086 TCCTTTGATCTTGTAAAGATGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142  
Db 897 ACATTTACATGCTCTCTCGTTCAGGTGATGATCAAGTTGTTGATTTGTTGCTTGAATA 956  
QY 2143 TGCAGCTTTGCTAGATGCTGCCAAGAGGCTTGTGTTAGCCAGAGTGAAGAAGTTGCTTC 2202  
Db 957 AGGAGCTCCAATCAGTCTGAAGACAAAAAATGTTTGGCTCCCTTACATATGCGACACA 1016  
QY 2203 TCTGATAATGTAANTTGGCGGATACCCAAAGCAGACATTCACAGCTTTTACATTTAGC 2262  
Db 1017 GTGGATGATGTTACTGTTGACTATCTC-----ACTCCTTTCATGTGGC 1061  
QY 2263 AGCTGTTTATAATAATTAAAGATGTCAGAGTATTTGTTTACACACGAGCTGATGAA 2322  
Db 1062 TGCTCATGCGGACATGTCGCTGCGTAACTTTTGTGGATCGTAATCGTACCCGAA 1121  
QY 2323 TGCCCAAGACAAAGGAGGACTTATCTTTACATAATGACAGCATCTTACGGGATGTAGA 2382  
Db 1122 TGCTGAGCTCTCAATGGCTTTCACACCGTGCATATCGCTTGCAAAAAAATCGCATTA 1181  
QY 2383 TGTAGAGCTCTACTAATAAGTATATGATGTGTCAATGTCACGCAAAATGGGCTTT 2442  
Db 1182 AATTGCGAACTGCTACTGAAATACACCGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241  
QY 2443 CACACCTTTGACGAGAGCCCAAGGAGCAACACAGCTTTGTGCTTTGTTGTAGC 2502  
Db 1242 CTCACCGCTGCATGTCGCTGCTTTATGGTGCTATAAACAATTGTCATCTATTACTACA 1301  
QY 2503 CCATGAGCTGACCCGACTTTAAAATCAGGAAGGACAAACCTTTTAGATTTAGTTTC 2562  
Db 1302 ACAAGTGCTAATGCAATGTGGCTACAGTACGCGGTGAACAGCCCTCTTATTAGCTGC 1361  
QY 2563 A 2563  
Db 1362 A 1362

## RESULT 12

US-09-065-474-c/c

; Sequence 36, Application US/09065474

; Patent No. 6063599

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Blehm, E. Scot

; TITLE OF INVENTION: DIFOPILARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 171

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESSEE: Heska Corporation

; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: Wordperfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/065,474  
; FILING DATE: 24-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verser, Carol Talkington  
; REGISTRATION NUMBER: 37,459  
; REFERENCE/DOCKET NUMBER: HW-5-C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 970/493-7272  
; TELEFAX: 970/484-9505  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5235 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-09-065-474-36

Query Match 2.4%; Score 93; DB 3; Length 5235;

Best Local Similarity 45.0%; Pred. No. 1.8e-16;

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1373 GGAGCAACATCAATGAAAGCACTAAAGATTTCTGACTCTCTGCGACGTTGATCTGAG 1432  
Db 5058 GGACCGATATCAACACATGCAATGGAATGGCTTAAATGCAATGCTGGCTCCAAA 4999  
QY 1433 AAAGCTCAVAAATGATGTTGAAGTAGTGTAACATGAACCAAGGTTAATGCTCTG 1492  
Db 4998 GAAGTCAATCATGAAGTGTCCCGGAACCTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 4939  
QY 1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTGTCTATCTACAAACCTGC 1552  
Db 4938 ACATGAAAGGGTAACACAGGTTTACATATAGCATCATTGGCAGGACAAGAATTAATCGTC 4879  
QY 1553 CGCCTACTCTGAGCTATGGTGTGATCCTTAACATATATATCCCTTCAGGCGCTTACTGCT 1612  
Db 4878 ACAGTACTTGTGAAATGGTGTAAATGTTACGTACAATCACTAAACGGTTTTTACACCA 4819  
QY 1613 TTACAGATGGGAAATGAAATGTACAGCAACTCTCCACAGAGGTATCTCATTTAGGTAAT 1672  
Db 4818 CTTTACATGGCTGCCAAGAAAATCAGAAATCTGTTGACGCTATCTTCTTGGCCACAAT 4759  
QY 1673 TCAGAGGACAGACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1725  
Db 4758 GCAATCAAGCTTTAAGTACAGAAAGACGGTTTTACGCCACTGCGCTTGGCTTGAACAA 4699  
QY 1726 TGTAAAAAAGCTGTGTACTGTTCCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTC 1785  
Db 4698 GGTACAGATCGTGTGCTGCTGTTTGTGTTGAAAATGACACGCGGGAAGTG-CGCTT 4640  
QY 1786 TACACCCTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTGTACA 1845  
Db 4639 GCCAGCACTGCATATTTGCTGCTAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580  
QY 1846 GCATGAGCTGATGTGCATGCTTAAAGATAAAGGAGCGCTTGTACCTTTGCACAATCATG 1905  
Db 4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520  
QY 1906 TTCTTATGACATTAAGTGTGCAACACTTCTTGTAAACATGAGCAGTACTTAAATGT 1965  
Db 4519 TCACATGGAATGAGAACGTAGCACTGCTACTCGAAAAGGAGCAATGTGAATTA 4460





Db 4219 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCATGTGC 4175  
QY 2263 AGTGGTGTATAAATTTAGAAAGTTGACAGAGTATTTGTACAAACGAGGCGTGTGAA 2322  
Db 4174 TGCTCATTCGGGACATGTCGCTGCGTAAACTTTTGTGATCGTAATGCTGACCCGAA 4115  
QY 2323 TGCCCAAGCAAAAGGAGGACTTATTCCTTTACATAATGACAGATCTTACGGGCAATGTA 2382  
Db 4114 TGCTGAGCTCTCAATGCGTTTACACCGCTGCATATCGCTTGCAAAAAAATCGCATTA 4055  
QY 2383 TGTAGCAGCTCTACTAATAAAGTATAATGATGCTGTCAATGCGCAGGACAAATGGCTTT 2442  
Db 4054 AATTGTGCAACTGCTACTGAATACCAAGCTGCAATCGAAGCACTACTGAATCGGCT 3995  
QY 2443 CACACCTTTGCGAAGAGCAGCCCAAAAGGAGGACACAGCTTTGTGCTTGTGCTAGC 2502  
Db 3994 CTCACGCTGCATGTGCTGCTTTTATGGGTGCTATAAACATTTGATCTATTTACTACA 3935  
QY 2503 CCATGAGCTGACCCGACTCTTTAAAAATCAGGAAGGACAAACACCTTTTAGATTAGTTTC 2562  
Db 3934 ACAAGTGTATTCAGATGTGGCTACAGTACGCGGTGAACGCTCTTCATTTAGCTGC 3875  
QY 2563 A 2563  
Db 3874 A 3874

## RESULT 15

US-09-031-485-32  
; Sequence 32, Application US/09031485  
; Patent No. 5824306  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Liang  
; APPLICANT: Blehm, E. Scot  
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; ADDRESS: Heskia Corporation  
; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: WordPerfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/031,485  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/847,429  
; FILING DATE: 24-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verser, Carol Talkington  
; REGISTRATION NUMBER: 37,459  
; REFERENCE/DOCKET NUMBER: HW-5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 970/493-7272  
; TELEFAX: 970/484-9505  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5503 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 51..5285  
; US-09-031-485-32  
  
Query Match 2.4%; Score 93; DB 1; Length 5503;  
Best Local Similarity 45.0%; Pred. No. 1.8e-16;  
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;  
  
QY 1373 GGAGCAACATCAATGAAAGACTTAAAGATTTCTTGACTCTCTGCACGTGGCATCTGAG 1432  
Db 228 GGCACCGATATACACACATGCAATGCGAATGGCCTTAATGCATTTGCATCTGGCCTCCAAA 287  
QY 1433 AAAGCTCATATGATGTTGTTGAAGTAGTGTGAACATGAAGCAAAAGGTTAATGCTCTG 1492  
Db 288 GAAGTCAATCAATGAATGCTGCGGCACTTCTGAAAGAAAGACGATGTTGATGCTGCC 347  
QY 1493 GATAATCTTGCTCAGACTTCTTACACAGAGCTGCATATTTGTGTGCTATCTACAACTGC 1552  
Db 348 ACTAGAAAGGCTAACACAGCGTTTACATATAGCATCAATTTGGCAGGACAAAGAACTAATCGTC 407  
QY 1553 GCGCTACTCTCTGAGCTATGGGTGTGATCCTAACATTTATCCCTTCAGGGCTTTACTGCT 1612  
Db 408 ACAGTACTTGTGAAAAATGGTCTAATGTTAACTGATACATCACTAAACGGTTTTACACCA 467  
QY 1613 TTACAGATGGGAATGAAAAATGTACAGCAACTCTCTCAAGAGGGTATCTCTATTAGTAAT 1672  
Db 468 CTTTACATGGCTGCACAAGAAAAATCAGCAATCTGTTGTACGCTATCTTCTTGCCCCACAT 527  
QY 1673 TCAGAGGCGACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGT-----CGAAAC 1725  
Db 528 GCCAATCAAGCTTTAAGTACAGAAAGACGCTTTTACGCCACTGGCAGTTGCTTGCACAA 587  
QY 1726 TGTAAAAAACTGTCTACTGTTTCAGAGTGTCAACTGCGAGACACATTTGAGGGCGTCAGTC 1785  
Db 588 GGTACAGATCGTGTGCTGCTGTTTGTCTGAAAATGACGCGCGGAAAGTG-CGCTT 646  
QY 1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGCTACA 1845  
Db 647 GCCAGCACTGCATATTTGCTGCTAAAAAGATGATACGAAAGCAGCTAGCTATTACTTCA 706  
QY 1846 GCATGAGCTGATGTCATGCTAAAGATAAGAGGAGGCGCTTGTACTTTGCAATGCAATG 1905  
Db 707 AAATGAGCAATACTCGGATGTGCTTTCGAAAAGGCGCTTTTACTCCGCTTCATATCGCGC 766  
QY 1906 TCTTTATGACATTTAAGTGTGAGAACTCTTGTGTTAAACATGAGCAGTAGTTAATGT 1965  
Db 767 TCACATATGGAATGAGAACGCTAGCACAACTGCTACTCGAAAAGGAGGCAATGTGAATTA 826  
QY 1966 AGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAT 2025  
Db 827 CCAACGAGACATAACATAAGTCCGTTACACGTTTGCACAAAAATGGGTCGTCACAAACAT 886  
QY 2026 TTGCAAACTTCTGCTCCAGCATGTTGCGAGACCCCTACCAAAAAAAGAGGATGAAATAC 2085  
Db 887 GGTTCGTTTATTTGTTGCTTCATGGGCGCTAATTTGACTGTGCGCACAGCTGATTTACTAAC 946  
QY 2086 TCCTTTGATCTTGTAAAGATGGAGATACAGAT---ATTCAAGATCTGTTAGGGAGA 2142  
Db 947 ACCATTACATCTGCTCTCGTTTACGTCATGATCAAGTTGTTGATTTGTTGCTTGAATA 1006  
QY 2143 TGCAGCTTTGCTTAGATGCTGCGCAAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTTC 2202  
Db 1007 AGGAGCTCCATCAGTCTAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGACA 1066  
QY 2203 TCCTGATATGTAATGTCGCGGATACCCAGGAGACATTCACACCTTTTACATTTAGC 2262  
Db 1067 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTCTGTTGCTG 1111  
QY 2263 AGCTGTTTATAATAATTTAGAGTTGACAGATTTTGTGTACAAACGCGGCTGATGTAA 2322  
Db 1112 TGCTCATTTGGGACATGTCGCTAAACTTTTGTGGATCGTAATGCTGACCCGAA 1171

QY	2323	TGCCAAGACAAGAGGAGCTTATTCCTTTACATAATGACAGCATCTTACGGGCATCTAGA	2382
Db	1172	TGCTGAGCTTCATGGCTTACACCGCTGCATATGCTTGGAAAAAATCGCATTA	1231
QY	2383	TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGACAAAATGGCTTT	2442
Db	1232	AATTGTGCACTTGCTTCTGAATACCCAGCTGCAATCGAAGCAACTACTGAATCCGGTCT	1291
QY	2443	CACACCTTTTGACAGCAAGCAGCCCAAAAGGGACGACAGCTTTGTGCTTTGTGTGAC	2502
Db	1292	CTCAGCGTGCATGTCGCTGCTTTATGGTGCATTAACATTTGCATCTATTACTACA	1351
QY	2503	CCATGGAGTCACCGACTCTTAAANAATCAGGAAGGACAACACCTTTAGATTATGTTTC	2562
Db	1352	ACAAGTGCTAATGCAGATGTGGCTACAGTACGGGGTGAACGCCCTCTCATTTAGCTGC	1411
QY	2563	A 2563	
Db	1412	A 1412	

Search completed: February 11, 2003, 18:30:39  
Job time : 179.212 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	3297	86.8	3400	10	US-09-509-1966A-1		Sequence 1, Appli
2	1813.4	47.8	4134	10	US-09-841-835-1		Sequence 1, Appli
3	1449.6	38.2	4491	10	US-09-841-835-7		Sequence 7, Appli
4	1297.8	34.2	4657	10	US-09-841-835-9		Sequence 9, Appli
5	994	26.2	2409	9	US-09-964-899-40		Sequence 40, Appli
6	467.6	12.3	1069	10	US-09-833-381-841		Sequence 841, Appli
7	461	12.1	465	10	US-09-833-381-1153		Sequence 1153, Ap
8	333	8.5	353	9	US-10-040-739-1179		Sequence 1179, Ap
9	121.2	3.2	523	10	US-09-841-835-12		Sequence 12, Appli
10	117.6	3.1	304	10	US-09-783-590-10862		Sequence 10862, A
11	81.6	2.1	466	10	US-09-833-381-842		Sequence 842, Appl
12	80.8	2.1	5175	9	US-09-964-899-42		Sequence 42, Appli
13	77	2.0	1299	10	US-09-908-711-12		Sequence 12, Appli
14	74.2	2.0	2505	10	US-09-947-199-3		Sequence 3, Appli
15	74.2	2.0	3025	10	US-09-947-199-1		Sequence 1, Appli
16	73.6	1.9	948	9	US-09-938-842A-2359		Sequence 2359, Ap
17	73.2	1.9	2271	10	US-09-835-7888A-2		Sequence 2, Appli
18	72.4	1.9	744	9	US-09-938-842A-805		Sequence 805, Appl
19	69.6	1.8	357	10	US-09-841-835-11		Sequence 11, Appli



Db 181 CGAAATACAGATGGAAGGACAGCATTTGGATTATTAGCAGATCCATCTGCCAAAGCAGTCTT 240  
QY 722 ACTGGTGAATATAGAAGATGAACCTCTTAGAAAGTGCAGGAGTGGCAATGAAGAAAA 781  
Db 241 ACTGGTGAATATAGAAGATGAACCTCTTAGAAAGTGCAGGAGTGGCAATGAAGAAAA 300  
QY 782 ATGATGGCTCTACTCACACCATTAATGTCACTGCCACGAAGTGAATGATGCGCAGAAAGTCA 841  
Db 301 ATGATGGCTCTACTCACACCATTAATGTCACTGCCACGAAGTGAATGATGCGCAGAAAGTCA 360  
QY 842 ACTCCATTACATTTGGCAGCAGATATACAGAGTAAAGATTGTACAGCTGTTACTGCAA 901  
Db 361 ACTCCATTACATTTGGCAGCAGATATACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420  
QY 902 CATGGAGCTGATGCCATGCTAAAGATAAAGTGAATCTGGTACCATTAACAAATGCCTGT 961  
Db 421 CATGGAGCTGATGCCATGCTAAAGATAAAGTGAATCTGGTACCATTAACAAATGCCTGT 480  
QY 962 TCTTATGGTCAATTAAGATTAAGTGAATCTGGTCAAGCATGGTGCCTGTGTPAATGCA 1021  
Db 481 TCTTATGGTCAATTAAGATTAAGTGAATCTGGTCAAGCATGGTGCCTGTGTPAATGCA 540  
QY 1022 ATGGACTTGTGGCAATTCACCTCTCTCAATGAGCAGCTCTTAAGACAGGGTTGAAGTA 1081  
Db 541 ATGGACTTGTGGCAATTCACCTCTCTCAATGAGCAGCTCTTAAGACAGGGTTGAAGTA 600  
QY 1082 TGTCTCTCTCTTAAAGTATGGTGAGACCCCAACACTGCTCAATTTGTACAAATAAAGT 1141  
Db 601 TGTCTCTCTCTTAAAGTATGGTGAGACCCCAACACTGCTCAATTTGTAAAGATAAAGT 660  
QY 1142 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 1201  
Db 661 GCTATAGACTTGGCTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGC 720  
QY 1202 CACTCGTCTGCAAGTGCACAGAGCTGATGTTACTCGAATCAAAAACATCTCTCT 1261  
Db 721 CACTCGTCTGCAAGTGCACAGAGCTGATGTTACTCGAATCAAAAACATCTCTCT 780  
QY 1262 CTGGAATGGTGAATTTCAAGCATCTCAACACATGAACAGCATTGATTTGCTGCT 1321  
Db 781 CTGGAATGGTGAATTTCAAGCATCTCAACACATGAACAGCATTGATTTGCTGCT 840  
QY 1322 GCATCTCCATATCCAAAAGAAAGCAAAATATGTGAATGTTGTAAGAAAGAGCAAAAC 1381  
Db 841 GCATCTCCATATCCAAAAGAAAGCAAAATATGTGAATGTTGTAAGAAAGAGCAAAAC 900  
QY 1382 ATCAATGAAAGACTAAAGAAATTTCTTGACTCTCTGACGTGGCATCTGAGAAAGCTCAT 1441  
Db 901 ATCAATGAAAGACTAAAGAAATTTCTTGACTCTCTGACGTGGCATCTGAGAAAGCTCAT 960  
QY 1442 AATGATGTTTGAAGTAGTGGTGAACATGAAGCAAAAGSTTAATGCTCTGGATAATCTT 1501  
Db 961 AATGATGTTTGAAGTAGTGGTGAACATGAAGCAAAAGSTTAATGCTCTGGATAATCTT 1020  
QY 1502 GGTACACTTCTACACAGAGTGCATATTTGGTGCATCTCAAAACCTGCCGCCCTACTC 1561  
Db 1021 GGTACACTTCTACACAGAGTGCATATTTGGTGCATCTCAAAACCTGCCGCCCTACTC 1080  
QY 1562 CTGAGCTATGGGTGATCCTTAACATTAATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621  
Db 1081 CTGAGCTATGGGTGATCCTTAACATTAATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140  
QY 1622 GGAATGAAATGTACAGCAACTCTCCAAAGGGTATCTCATTTAGGTAAATTCAGAGGCA 1681  
Db 1141 GGAATGAAATGTACAGCAACTCTCCAAAGGGTATCTCATTTAGGTAAATTCAGAGGCA 1200  
QY 1682 GACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACCTGTGT 1741  
Db 1201 GACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACCTGTGT 1260  
QY 1742 ACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCCTTCATTTT 1801  
Db 1261 ACTGTTTCAGAGTCTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCCTTCATTTT 1320

QY 1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGTACAGCATGGAGCTGATGTG 1861  
Db 1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGTACAGCATGGAGCTGATGTG 1380  
QY 1862 CATGCTAAAGATAAAGAGGCCCTTGACCTTTGCACAAATGCATGTTCTTTATGACATTAAT 1921  
Db 1381 CATGCTAAAGATAAAGAGGCCCTTGACCTTTGCACAAATGCATGTTCTTTACGGACATTAAT 1440  
QY 1922 GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTAAATGTAGCTGATTTATGAAA 1981  
Db 1441 GAAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTAAATGTAGCTGATTTATGAAA 1500  
QY 1982 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTTCTGCTC 2041  
Db 1501 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTTCTGCTC 1560  
QY 2042 CAGCATGGTGCAGACCTTACCACCAAAACAGGAGTGGAAATACTCCTTTTGGATCTTTGTT 2101  
Db 1561 CAGCATGGTGCAGACCTTACCACCAAAACAGGAGTGGAAATACTCCTTTTGGATCTTTGTT 1620  
QY 2102 AAAGATGGAGATACAGATATTCAGAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 2161  
Db 1621 AAAGATGGAGATACAGATATTCAGAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 1680  
QY 2162 GCCAAGAAGGGTGTGTAGCCAGAGTGAAGAAGTGTCTTCTCTGATATGTAATTTGCT 2221  
Db 1681 GCCAAGAAGGGTGTGTAGCCAGAGTGAAGAAGTGTCTTCTCTGATATGTAATTTGCT 1740  
QY 2222 CGGGATACCCAGGACAGATTCACACCTTTTACATTTAGCAGCTGGTATATATTAATTA 2281  
Db 1741 CGGGATACCCAGGACAGATTCACACCTTTTACATTTAGCAGCTGGTATATATTAATTA 1800  
QY 2282 GAAGTTGCAGAGTATTTGTACAAACAGGAGCTGATGTAATGCCCAAGCAAGAGGGA 2341  
Db 1801 GAAGTTGCAGAGTATTTGTACAAACAGGAGCTGATGTAATGCCCAAGCAAGAGGGA 1860  
QY 2342 CTTTATTCCTTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2401  
Db 1861 CTTTATTCCTTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920  
QY 2402 AAGTATATGCATGTGCTAATGCCAGGACAAATGGCTTTTCACACCTTTGACAGGA 2461  
Db 1921 AAGTATATGCATGTGCTAATGCCAGGACAAATGGCTTTTCACACCTTTGACAGGA 1980  
QY 2462 GCCCAAAAGGAGCAACACAGCTTTTGTGCTTGTAGCCCATGGAGCTGACCCGACT 2521  
Db 1981 GCCCAAAAGGAGCAACACAGCTTTTGTGCTTGTAGCCCATGGAGCTGACCCGACT 2040  
QY 2522 CTTTAAAAATCAGGAAGCAAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTACGGCT 2581  
Db 2041 CTTTAAAAATCAGGAAGCAAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTACGGCT 2100  
QY 2582 CTTTCTCAGCAGCCCATGCCCCCATCTGCTCTGCCCTCTTGTGTACAGCCTCAAGTGCTC 2641  
Db 2101 CTTTCTCAGCAGCCCATGCCCCCATCTGCTCTGCCCTCTTGTGTACAGCCTCAAGTGCTC 2160  
QY 2642 AATGTTGTGAGAGCCACGAGCCACTGCAGATGCTCTCTCTTCAGCTCCATCTAGCCCA 2701  
Db 2161 AATGTTGTGAGAGCCACGAGCCACTGCAGATGCTCTCTCTTCAGCTCCATCTAGCCCA 2220  
QY 2702 TCAAGCCTTTCTCAGCCAGCAGCTTTGACAACCTTATCTGGGAGTTTTTCAGAACTGTCT 2761  
Db 2221 TCAAGCCTTTCTCAGCCAGCAGCTTTGACAACCTTATCTGGGAGTTTTTCAGAACTGTCT 2280  
QY 2762 TCAGTAGTTTAGTTCAAGTGAACAGAGGTGCTTCCAGTTTGGAGAAAAGAGGAGTTCCA 2821  
Db 2281 TCAGTAGTTTAGTTCAAGTGAACAGAGGTGCTTCCAGTTTGGAGAAAAGAGGAGTTCCA 2340  
QY 2822 GGAAGTAGTTTAGCTAACTCAATTCGTAAGGAATCTTGGACTTCAGCACCTCAATGGAT 2881  
Db 2341 GGAAGTAGTTTAGCTAACTCAATTCGTAAGGAATCTTGGACTTCAGCACCTCAATGGAT 2400

Qy	2382	ATATTTGAGAGAGACAGATCACTTTGGGATGTTATTAGTTGAGATGGGGCACAAGGAGCTG	2391
Db	2401	ATATTTTGAGAGAGAACAGATCACTTTGGGATGTTATTAGTTGAGATGGGGCACAAGGAGCTG	2460
Qy	2942	AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACACTAATTTAAGGAGTCTCGAGAGA	3001
Db	2461	AAGGAGATTGGGAATCAATGCTTATGGACATAGGCACAACACTAATTTAAGGAGTCTCGAGAGA	2520
Qy	3002	CTTATCTCCGGACAACAAGGCTCTTAACCCATATTTTAACTTTGAACACCTCTGGTAGTGGGA	3061
Db	2521	CTTATCTCCGGACAACAAGGCTCTTAAGCCATATTTTAACTTTGAACACCTCTGGTAGTGGGA	2580
Qy	3062	ACAAATCTTATAGATCTGTCTCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG	3121
Db	2581	ACAAATCTTATAGATCTGTCTCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG	2640
Qy	3122	CAAAGTACAGTTCCGAGAGCACAGAGATGGAGTCAATGAGGTGGAACTTTCAACAGATAC	3181
Db	2641	CAAAGTACAGTTCCGAGAGCACAGAGATGGAGTCAATGAGGTGGAACTTTCAACAGATAC	2700
Qy	3182	AATATTCTCAAGATTTCAGAAAGTTTGTAAACAAAGAACTATGGGAAAAGATACACTCACCGG	3241
Db	2701	AATATTCTCAAGATTTCAGAAAGTTTGTAAACAAAGAACTATGGGAAAAGATACACTCACCGG	2760
Qy	3242	AGAAAAGAGTTTCTGAGAGAAACCAACCAACCATGCGAATGAACGAATGCTATTTTCATGGG	3301
Db	2761	AGAAAAGAGTTTCTGAGAGAAACCAACCAACCATGCGAATGAACGAATGCTATTTTCATGGG	2820
Qy	3302	TCCTCCCTTTGTGAATGCAATATTCACAAAAGGCTTTGATGAAAGGCATGGCTACATAGGT	3361
Db	2821	TCCTCCCTTTGTGAATGCAATATTCACAAAAGGCTTTGATGAAAGGCATGGCTACATAGGT	2880
Qy	3362	GGTATGTTGGAGCTGGCATTTATTTTGTCTGAAAACCTCTCCAAAAGCAATCAATATGTA	3421
Db	2881	GGTATGTTGGAGCTGGCATTTATTTTGTCTGAAAACCTCTCCAAAAGCAATCAATATGTA	2940
Qy	3422	TATGGAAATGGAGGAGGTACTGGGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC	3481
Db	2941	TATGGAAATGGAGGAGGTACTGGGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC	3000
Qy	3482	CACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTTCAGTGCA	3541
Db	3001	CACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTTCAGTGCA	3060
Qy	3542	ATGAAAATGGCACATTTCTCCTCAGGTCAATCACTCAGTCACTGCTAGGCCAGCTGTAAT	3601
Db	3061	ATGAAAATGGCACATTTCTCCTCAGGTCAATCACTCAGTCACTGCTAGGCCAGCTGTAAT	3120
Qy	3602	GGCCTAGCATTTAGCTGAATATGTTATTTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA	3661
Db	3121	GGCCTAGCATTTAGCTGAATATGTTATTTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA	3180
Qy	3662	ATTACTTACCAGATTTATGAGGCTTGAAGGTATGTTGCGATGGATAAATAGTATTTTAAGA	3721
Db	3181	ATTACTTACCAGATTTATGAGGCTTGAAGGTATGTTGCGATGGATAAATAGTATTTTAAGA	3240
Qy	3722	AACATAATCCACTGAACTTAAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTACTCCCT	3781
Db	3241	AACATAATCCACTGAACTTAAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTACTCCCT	3300
Qy	3782	TGCTGAAAAAAA 3794	
Db	3301	TGCTGAAAAAAA 3313	

## RESULT 2

```

US-09-841-835-1
; Sequence 1, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

```

Db 1077 CTAATGTGAATGCCATGCAAGTGTGGCGAAGATCGCACTCCTTTACATCTAGCAGCG 11136  
QY 863 GGATATAACAGAGTAAGATTGTACAGCTGTCTACTGCAACATGGAGCTGATGTCCATGCT 922  
Db 1137 GGCTACAACAGAGTTGGAATAGTTCAGCTTCTCTTCAGCATGGTGTGATGTTCAATGCA 11196  
QY 923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCCTGTCTTATGGTCATTTAAGTA 982  
Db 1197 AAAGACAAGGTGGACTGTGGCTCTTCATAATGCATGTTTCATATGGACATTTAAGTC 1256  
QY 983 ACTGAACCTTTTGGTCAAGCATGGTGGCTGTGTAATGCAATGCAATGGCTGTGGCAATTCAT 1042  
Db 1257 ACAGAATCTGTACTAAAGCATGAGCTTGTGTTAATGCCATGATCTCTGGCAGTTTACT 1316  
QY 1043 CTTCTTCATGAGCGCTCTTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTTAAGTTAT 1102  
Db 1317 CCACTGCACGAGGCTGCTTCCAAGACCGGTGTAAGTCTGTCTTTGTTTACTTAGCCAT 1376  
QY 1103 GGTGCAGACCCAACTGCTCAATTTGTCAAAATTAAGTGTCTATAGCATTTGGTCCACCA 11162  
Db 1377 GGCCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTCTGTGGATATGGCTCCAAC 1436  
QY 1163 CCACAGTTAAACAAGATTAGCATATGAATTTAAAGGCCACTGCTGTGCAAGTGC 1222  
Db 1437 CCGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCAITTTTACTACAAAGC 1496  
QY 1223 CGAAGCTGATGTTACTCGAAATCAAAAACATCTCTCTCGAAATGGTGAATTTCAAG 1282  
Db 1497 AGAAGCAGACCTTAGCTAAGTTAAAAAACACTCGCTCGAATCATTAATTTCAA 1556  
QY 1283 CATCTCAAAACATGAACAGCATTTGCAATTTGTGCTGCATCTCCATATCCCAAGA 1342  
Db 1557 CAACCGCAGTCTCATGAACAGCACTGCACCTGTCTGTGGCTCTCTGCATCCCAACGT 1616  
QY 1343 AAGCAATATGTGAACCTTTGCTTAAGAAAAGGCAACATCAATGAAGAGCTTAAGAA 1402  
Db 1617 AAACAAGTGACAGATTTGTACTTTAAGAAAAGGAGCAATGTTAATGAAGAAAATAAAGAT 1676  
QY 1403 TTTCTGACTCTCTGCACTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTG 1462  
Db 1677 TTTATGACTCCCTGCACTGTTGAGCGGAAAGACCCATATGATGTCATGGAAGTTCTG 1736  
QY 1463 GTGAACATGAAGCAAGGTTAATGCTCTGGATAATCTGGTCAGACTTCTCTACACAGA 1522  
Db 1737 CATAAAGCATGGCCAAAGATGATGCACTGGACACCCCTGGTCAGACTGCTTTGCATAGA 1796  
QY 1523 GCTGCATATTTGCTCATCTACAAACCTCGCGCTACTCTGAGCTATGGGTGTGATCCT 1582  
Db 1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCGCCTCTGCTGAGTTACGGCTCTGACCCC 1856  
QY 1583 AACATTTATCCCTTACAGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAA 1642  
Db 1857 TCCATCATCTCCTTACAAGGCTTACAGCAGACAGATGGCAATGAAGCAGTGCAGCAG 1916  
QY 1643 CTCTCCAAGAGGTATCTCATTTAGTAAATTCAGAGCCACACAGACAATTTGCTGGAAGCT 1702  
Db 1917 ATTCTGAGTGAGAGTACACCTATACGTTACTTCTGATGTTGATTCAGACTCTTAGAGGCA 1976  
QY 1703 GCAAGGCTGGAGATGCGAAACTGAAAAAACTGTGTACTGTTACAGAGTGTCAACTGC 1762  
Db 1977 TCTAAGCTGGAGCTTGGAACTGTGAACCACTTTGACGACTCTCAAAATGTGAATGT 2036  
QY 1763 AGAGACATTAAGGGGCTCAGTCTACACCCTTCATTTTGCAGCTGGGTATAACAGAGTG 1822  
Db 2037 AGAGACTTAGAGGGCCGGCATTCACGCCCTTACACTTCGCAAGGCTCAACCCGGTG 2096  
QY 1823 TCCGTGTTGAATATCTGCTACAGCATGGAGCTCATGTGATGCTAAAGATAAAGGAGGC 1882  
Db 2097 TCTGTTAGAGTACCTGCTACACACCGGTGCGGATGTCCATGCGCAAGACAAAGGGTGGC 2156  
QY 1883 CTTGTACTTTGCAACATGATGTTCTTATGGACATTTATGAAGTTGCAAGACTTCTTCTT 1942  
Db 2157 TTGGTCCCTTTCATATGCTGTTTCATATCGACACTATGAGGGTGGCTGAGCTTTTAGTA 2216

QY 1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAATTTACACCTTTTACATGAAGCA 2002  
Db 2217 AGCATGGGGCTTCTGCTCAATGTGGCGGACTTATGGNAATTTACCCTCTCCATGAAGCA 2276  
QY 2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCAGCATGGTGCAGACCCCTACC 2062  
Db 2277 GCAGCTAAAGGAAAATGAAATCTGCAAGCTCCTTTTAAACATGGAGCAGATCCAAC 2336  
QY 2063 AAAAAACAGGATGAAATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATT 2122  
Db 2337 AAAAAACAGATGAAATACACCTTTGGATTTGTAAGGAAGGAGACACAGATATT 2396  
QY 2123 CAAGATCTGTTAGGCGAGATGAGCTTTTGTAGTCTGCCAAGAGAGGTTGTTTAGCC 2182  
Db 2397 CAGGACTTACTGAAGGGGATGCTGCTTTGTTGATGCTGCCAAGAGGGCTGCCCTGGCA 2456  
QY 2183 AGAGTGAAGAAAGTTGCTTCTCTGATAATGTAAATTTGCCGCGATACCCAAAGGAGACAT 2242  
Db 2457 AGAGTGCAGAAAGCTGTACCCCGAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAAT 2516  
QY 2243 TCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTCTTA 2302  
Db 2517 TCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTCTA 2576  
QY 2303 CAACACGGAGCTGATGTGAATGCCCAAGACAAGAGGAGCTTATTCCTTTTACATAATGCA 2362  
Db 2577 GAGCATGGAGCTGATGTTAATGCCAGCAAGGGTGTTTAAATTCCTCTTCAATATGCG 2636  
QY 2363 GCATCTTACGGGATGTAGATGTAGCAGCTCTACTATAAAGTATATATGATGTGTCAAT 2422  
Db 2637 GCATCTTATGGGATGTGACATAGCGCTTTATTTGAATAATACAAACAGCTGTGTAAAT 2696  
QY 2423 GCCACGGACAATATGGGCTTTCAACACCTTTGCAGCAAGCAGCCCAAGAGGACGAACACAG 2482  
Db 2697 GCAACAGATAAGTGGCGTTTACTCCCTCCATGAACGCCAGCAAGAGGAGCCGAG 2756  
QY 2483 CTTTGTGCTTTGTTGTAGCCCATGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAA 2542  
Db 2757 CTGTGCGCCCTCTCTCTAGCGCATGGTCAGACCCCAACCATGAAGAACCAAGAGGCGCAG 2816  
QY 2543 ACACCTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCTCTCTGACAGCAGCCATGCC 2602  
Db 2817 ACGCTCTGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCC 2876  
QY 2603 CCATCTGCTCTGCCCTCTTGTATCAAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCCAAGGA 2662  
Db 2877 CCAGAGGCCCTTACTACTCTGTTTAAACCTCAGG-----CTACT 2915  
QY 2663 GCCACTGCAGATGCTCTCTCTTCTAGGTCCATCTAGCCCCATCAAGCCTTTCTGACAGCAGC 2722  
Db 2916 GTAGTGAGTGCCCTCTCTGATCTCACAGCATCCACCCCTCTCGCTCTGGCTGCCAGC 2975  
QY 2723 AGTCTTCACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGA 2782  
Db 2976 AGCATAGCAACCTCACTGGCCCTTTAGCAGAGTTGCCGTAGGAGGAGCCCTCAATGCA 3035  
QY 2783 ACAGAGGCTGCTTCCAGTTTGGAGAAAAG-----GAGGTTCCAGGAGTAGATTTTAGC 2836  
Db 3036 GGGATGCGCGCGGGAACACAGAAAGGAAGGAAGGAAAGTTGCTGTCTTGACATGAAT 3095  
QY 2837 ATAACTCAATTCGTAAGGAATCTTGGAGCTTGAGCACCTAATGGATATATTTTGAAGAGAA 2896  
Db 3096 ATCAGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTTGAACAGAA 3155  
QY 2897 CAGATCACTTTGGATGTTAGTTGAGATGGGCAAGAGAGCTGAAGGAGATTTGGAATC 2956  
Db 3156 CAGATTCACATAGATGTTGGCTGATATGGTTCATGAAGAGTTGAAGAATATAGGCATC 3215  
QY 2957 AATGCTTATGACATAGGCAACAACTAATTAAGAGAGTCGAGAGACTTATCTCGGACAA 3016  
Db 3216 AATGCATATGGCCACCGCCCAACAAATTAATCAAGAGAGTGAAGAAGACTCTTAGGTGGACAA 3275

QY 3017 CAAGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGAACAATTTCTTATAGAT 3076  
Db 3276 CAAGGCACCAATCTTATTTGACCTTTTCACTGTGTTAATCAGGGAACGATTTTCTGTGAT 3335  
QY 3077 CTGCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGATGCAAAAGTACAGTTTGA 3136  
Db 3336 CTGCTCCAGAAGATAAGATATCACTGAGTGGGAAGAGATGCAAAAGTACTATTGCA 3395  
QY 3137 GAGCACAGATGAGGTCTATGAGGTGGAATCTTCAACAGATACATATTTCTCAAGATT 3196  
Db 3396 GAACACAGAGATGTTGATGCTGGCGGATCTTCAACAGATACATGATGATTCGAAAT 3455  
QY 3197 CAGAAGTTTGTACAGAAGAACTATGGGAAGATACACTCACCGGAGAAAGATTTCT 3256  
Db 3456 CAAAAGTTGTCAAGAAGAACTTGAAGGAGCGTCTGCCACCGACAGAGAGAGTGTCT 3515  
QY 3257 GAAGAAACACCAACCATGCGCAATGCAATGCAATGCTATTTTCATGGTCTCTTTTGAAT 3316  
Db 3516 GAGGAGATCAACCAATCAATGAGCGGATGTTTTCATGGTCTCTCTTTTCATTAAT 3575  
QY 3317 GCAATTATCCCAAGGCTTTGATGAAAGGATGCGGTACATAGTGTGATGTTTGGAGCT 3376  
Db 3576 GCCATTATCATAAAGGTTTGAAGCGCATGCGATACATAGAGGAATGTTGGGGCC 3635  
QY 3377 GGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATATATGAAATGGAGGA 3436  
Db 3636 GGGATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATATGAAATGGAGGA 3695  
QY 3437 GGTACTGGGTGCTCAGTTCACAAAGCAGATCTGTTACATTTGCCACAGGAGCTGCTC 3496  
Db 3696 GGAACAGCTGCTTACACAGAGCAGGTCATGCTATATATGTCACAGACAAATGCTC 3755  
QY 3497 TTTTGGCGGTAACTTTGGGAAAGTCTTTCTGCACTGAGTGCATGAAATGGCACAT 3556  
Db 3756 TTTCTGATAGTGACCTTTGGGAAATCTTTCTGCACTGAGTGCATGAAATGGCCCCAC 3815  
QY 3557 TCCTCTCCAGGTATCACTCAGTCTAGTGTAGGCCAGTGTAAATGGCTAGCATTAGCT 3616  
Db 3816 GCGCTCCAGGGCACCCTCAGTCTAGTGTAGGCCAGTGTAAATGGCTAGCATTAGCT 3875  
QY 3617 GAATATGTTTATACAGAGCAGAACAGGCTTATCTCTGAGTATTTATTAATCTACAGATT 3676  
Db 3876 GAATATGTCATCTACAGAGGAGAACAGGCTATCCACAGAGTATCTTATCACTTACCAGATC 3935  
QY 3677 ATGAGGCTGAAG 3689  
Db 3936 ATGAAGCCAGAAG 3948

## RESULT 3

US-09-841-835-7  
; Sequence 7, Application US/09841835  
; Patent No. US2002007695A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835

; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4491 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 6..2027  
; US-09-841-835-7  
  
Query Match 38.2%; Score 1449.6; DB 10; Length 4491;  
Best Local Similarity 65.3%; Pred. No. 0;  
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;  
  
QY 383 TCCTCTGCTTTATCATTAAGGTTTTGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442  
Db 657 TCTCCCTGCACCTTCGCTGCAGGTTTTGGAAGGAAGATGTTGTAGAACACTTACTACAG 716  
QY 443 AATGTGCAAGTGTCACCAAGCAGTCATGATGGGGCCCTTATTCCTCTTCATAATCATGCG 502  
Db 717 ATGGGTGCTAATGTCCAGCTCGTCATGATGAGGTCTCATCCCGCTTCATAATGCTGT 776  
QY 503 TCCTTTGGTCATGCTGAAGTAGTCAATCTCTTTTGGCAGATGGTGGAGACCCCAATGCT 562  
Db 777 TCCTTTGGCAGTGGTGGAGGTTGTGAGTCTGTTATTTGCCAAGGAGCTGATCCAAATGCC 836  
QY 563 CGAGATAATTGGAATTATATCTCTCTCATGAAGCTGCAATTTAAAGGAAAGATTCATGCTT 622  
Db 837 AGGGATACTGGAACATATACACCTCTGATGAGCTGCTATTAAGGGAAGATCGATGTG 896  
QY 623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACA 682  
Db 897 TGCATTGTGCTGCTGCAGCAGGAGCTGACCCAAACATTCGGAACACTGATGGGAATCA 956  
QY 683 GCATTGGATTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATTAAGAAAGAT 742  
Db 957 GCGCTGGACCTGGCAGATCCCTTCAGCAAAAGCTGTCTTACAGGTGAATACAAAGAAC 1016  
QY 743 GAACCTCTTAGAAGTGCCAGGATGGCAATGAAGAAAAAATGATGGCTCTTACTCACACCA 802  
Db 1017 GAACCTCTTAGAAGTGCTAGGAGTGGTAATGAAGAAAAAATGATGGCTTCTTACTGACTCT 1076  
QY 803 TTAATGTCAACTGCCACGCAAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862  
Db 1077 CTAATGTGAATTGCCATGCAAGTATGGCGAAAGTGCAGCTCTTTTACATCTAGCAGCG 1136  
QY 863 GGATATAACAGATTAAGATTGTACAGCTGTACTGCAACATGGAGCTGATGTCATGCT 922  
Db 1137 GGCTACAACAGATTGCAATAGTTCAGCTTCTTCTTACAGCATGGTGTGATGTTTCATGCA 1196  
QY 923 AAAGATAAGGTGATCTGGTACCATTACACAAATGCCTTCTTATGTCATTATGAAGTA 982  
Db 1197 AAAGACAAGGTGGACTGTGCTCTTCTCATTAATGCAATGATGAGACATTTATGAAGTC 1256  
QY 983 ACTGAATTTTGGTCAAGCATGGTGCCTTGTAAATGCAATGGAGCTTGTGGCAATTCAC 1042  
Db 1257 ACAGAACTGCTACTAAAGCATGGAGCTGTGTTAATGCCATGGATCTCTGCGAGTTTACT 1316

QY 1043 CTTCTTCATGAGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAAGTTAT 1102  
Db 1317 CCACTGCACAGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTGTACTTAGCCAT 1376  
QY 1103 GGTGCAGACCCAACTGCTCAATTTGTCAAAATAAAGTGCTATAGACTTTGGCTCCACA 1162  
Db 1377 GGGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAAT 1436  
QY 1163 CCACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGTGCAAGTGCA 1222  
Db 1437 CCGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTACTACAAAGGCC 1496  
QY 1223 CGAAGAGCTGATGTTACTCGAATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAG 1282  
Db 1497 AGAAGCAGCAGTTAGTAAAGTTAAAAAACAACACTCGCTCTGGAATCATTAATTTCAA 1556  
QY 1283 CATCTCTAAACACATGAACAGCAATTTGCAATTTGTGCTGCACTCTCCATATCCCAAGA 1342  
Db 1557 CAACCGAGTCTCATGAACAGCACTGCACTGTGCTGTGGCTCTCTGCACTCCCAACGT 1616  
QY 1343 AAGCAATATGTAAGTGTGCTTAAGAAAGAGCAACATCAATGAAGAGACTAAAGAA 1402  
Db 1617 AAACAAGTGACAGAAATTTGTACTTAGAAAAGAGCAAAATTTAATGAAGAAATAAAGAT 1676  
QY 1403 TTCTTGACTCTCTGCAGCTGGCATCTGAGAAAAGCTCATATGATGTTGTGAAGTAGTG 1462  
Db 1677 TTCATGCTCCCTGCTGATGTGAGCGGAAGAGCCCATTAATGATGTCATGGAAGTCTG 1736  
QY 1463 GTGAACATGAAGCAAGGTTAATGCTCTGATTAATCTTGGTCAGACTTCTCTACACAGA 1522  
Db 1737 CATAGCATGGCGCAAGATGAATGCACTGGACACCCCTTGGTCAGACTGCTTTCATAGA 1796  
QY 1523 GCTGCATATTTGGTGTATCTACAAACTGCGGCTACTCTGAGCTATGGTGTGATCCT 1582  
Db 1797 GCGGCCCTAGCAGGCCACCTGCAAGACCTGCGGCTCTGCTGAGTTACGGCTCTGACCCC 1856  
QY 1583 AACATTATATCCCTTACAGGCTTTACTGCTTTACAGATGGAAATGAAATGTACAGCAA 1642  
Db 1857 TCCATCATCTCTTACAGGCTTACAGAGCTTACAGAGCAGCAGATGGCAATGAAGCAGTCCAC 1916  
QY 1643 CT----- 1644  
Db 1917 ATTCTGAGTGTGATTACGGCTCTGACCCCTCCATCATCTCTTACAAAGCTTTCACAGCA 1976  
QY 1645 ----- 1644  
Db 1977 GCACAGATGGCAATGAAGCAGTGCACAGATTTCTGAGTGGTCATTCGTAGATAGTATC 2036  
QY 1645 ----- 1644  
Db 2037 ATTCTACTCAGCCTTAATGGTGTATTTGAGCGGGAAGATTTAGAAGGAAATCTATCCA 2096  
QY 1645 ----- 1644  
Db 2097 GCATGCTTCACTGTCAACATGAAGAGTACACCTTATACGTACTTCTGATGTTGATTATCG 2156  
QY 1645 ----- 1644  
Db 2157 ACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAACTGTGAAGCACTTTGCACCTCTCA 2216  
QY 1645 -----C 1645  
Db 2217 AATGTGAATTTAGAGACTTAGAGGGCGGCAATTCACAGCCCTTTACACTTCGCAGCAGG 2276  
QY 1646 CTCAAGAGGGTATCTCATTTAGTTAAATTCAGAGCCAGACAGCAATTTGCTGGAAGCTGCA 1705  
Db 2277 CTACAACAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCACTCT 2336  
QY 1706 AAGCTCGAGATGCGAACTGTAATAAAGCTGTACTGTTCAGAGTGTCAACTGCGAGA 1765  
Db 2337 AAAGCTGGAGACTTGGAACTGTGAAGCACTTTGCAGCTCTCAAAATGTGAATTTGTAGA 2396

QY 1766 GACATTGAAGGGCGTCAGTCTACACCACTTCAATTTGCAGCTGGGTATAACAGAGTGTCC 1825  
Db 2397 GACTTAGAGGGCGGCGATTCCACGCCCTTACACTTCCGACAGGCTACAAACCGCGTCT 2456  
QY 1826 GTGGTGAATATCTGTACAGCATGGAGCTGATGCTGCTAAAGATAAAGAGGCGCTT 1885  
Db 2457 GTTGTAGAGTACCTGTACACCGGTCGCGATGTCATGCCAAAGACAAAGGTGGCTG 2516  
QY 1886 GTACCTTTGCACAAATCATGCTTCTTATGGACATTTATGAAGTTTGCAGAACTTCTTGTAAA 1945  
Db 2517 GTCCCTTCATATGCTGCTTCAATGACACTATGAGTGCTGAGCTTTTAGTAAGG 2576  
QY 1946 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGA 2005  
Db 2577 CATGGGCTTCTCTCAATGTGGGACTTATGGAATTTTACCCCTCTCCATGAAGCAGA 2636  
QY 2006 GAAAAAGGAAAATATGAATTTGCAAACTTCTGCTCCAGCATGTCAGACCCCTACCAAA 2065  
Db 2637 GCTTAAAGGAAAGTATGAATCTCAAGCTCTTTTAAACATGGAGCAGATCCAACATAA 2696  
QY 2066 AAAAAAGGAGTGAATCTCTCTTGGATCTTGTAAAGATGGAGATACAGATATTCAA 2125  
Db 2697 AAGAACAGAGATGAATACACCTTTGGATTTGGTAAGGAAGAGACACAGATATTCA 2756  
QY 2126 GATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGA 2185  
Db 2757 GACTTACTGAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAGGGCTGCTGGCAAGA 2816  
QY 2186 GTCAAGAGTTGCTCTCTCTGATATGTAATTTGCCGCGTATCCCAAGCAGACATCA 2245  
Db 2817 GTCAAGAGTCTGTACCCAGAGATATCACTGTCAGACACACCCAGGCGAGAAATCA 2876  
QY 2246 ACACCTTTACATTTAGCAGCTGTTATAATTTTAAAGTTGCAGAGTATTTGTTACAA 2305  
Db 2877 ACCCTCTGCACCTGGCAGCAGCTATATAACCTGGAAGTAGTGTATCTCTAGAG 2936  
QY 2306 CACGGAGCTGATGTAATGCCAAGACAAAGAGGAGCTTATTCCTTTTACATAATGACGA 2365  
Db 2937 CATGGAGCTGATGTAATGCCAAGAGGGTGTAAATTCCTCTTCAATAATGCGGCA 2996  
QY 2366 TCTTACGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2425  
Db 2997 TCTTATGGGATGTTGATAGCGGCTTTATGATGATGATGATGATGATGATGATGATGATGAT 3056  
QY 2426 ACAGCAAAATGGGCTTTTACACCTTTTGCAGAGCAGCCCAAAAGGAGGACGACAGCTT 2485  
Db 3057 ACAGATAAGTGGGCTTTTACTCCCTCCATCAAGCAGCCCAAGAAAGGAGCAGCTG 3116  
QY 2486 TGTGCTTTGTTGCTAGCCCATGAGCTGACCCGACTCTTAAATATCAGGAGGACAAACA 2545  
Db 3117 TGGCCCTCTCTTACGCGCATGTCAGACCCACCATGAAAGAACAGGAGGCGCAGCG 3176  
QY 2546 CTTTAGATTTAGTTTACGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCTGCCCCCA 2605  
Db 3177 CCTGTGATCTGGCAACAGCTGAGATATCAGAGTTTGTGATGATGATGATGATGATGATGAT 3236  
QY 2606 TCTGCTTGCCTCTCTGTTTACAGCCTCAAGTGTCTAATGTTGTGAGAGCCCGAGGAGCC 2665  
Db 3237 GAGGCCCTTACTACTCTGTTTTAAACCTCAGG-----CTACTGTA 3275  
QY 2666 ACTGCAGATGCTCTCTTTCAGGTCCTATGAGCCCATCAAGCCCTTCTGACGCGCAGCT 2725  
Db 3276 GTGAGTGCCTCTGATGTCACAGCATCCACCCCTCTGCGCTCGGCTGCGCCAGCAGC 3335  
QY 2726 CTTCAACAATTTATCGGAGTGTTCAGAACTGCTTTCAGTAGTTAGTTTCAAGTGGAAACA 2785  
Db 3336 ATAGCAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGCGCTCCAATGCGAGG 3395  
QY 2786 GAGGCTCTTCCAGTTTGGAGAAAAAGAG-----GTTCCAGGAGTAGATTTTAGCAATA 2839  
Db 3396 GATGGCGCGCGGAACAGAAAGGAAGGAGGAGGAGTGTGCTGCTTGTGATGAATATC 3455  
QY 2840 ACTCAATTCGTAAAGAACTCTGGACTTTCAGCACCTAATGGATATATTTGAGAGAGAACAG 2899

Db 3456 AGCCAAATTTCTAAAGAGCTTGGCTTGAACACCTTCGGGATATCTTTGAAACAGACAG 3515  
QY 2900 ATCACTTTGGATATTTAGTTAGATGGGCGACAGAGCTGAAGAGATGGATCAAT 2959  
Db 3516 ATTACACTAGATGTTGGCTGATATGGTTCATGAAGAGTTGAAGAAATAGGCATCAAT 3575  
QY 2960 GCTTATGACATAGGCACAACTAATTAAGCAGTTCGAGAGACTTATCTCCGGACACAA 3019  
Db 3576 GCATATGGCCACCCACAAATTAATCAAGAGTAGAAGACTCTTAGTGGACACAA 3635  
QY 3020 GGTCTTAACCCCATTTAATCTTTGAACACCTCTCTGTAGTGGAACTTCTTTATAGATCTG 3079  
Db 3636 GCGACCAATCTTATTTGACTTTTCACTGCTGTAACTCAGGGAACGATTTTCTGCTGATCTT 3695  
QY 3080 TCTCTGATGATAAAGAGCTTTTCACTCTCTGGAGGAGATGCAANGTACAGTTTCGAGAG 3139  
Db 3696 GCTCAGAGATTAAGATATACGTCAGTGAAGAGAGATGCAAGTACTATTTCGAGAA 3755  
QY 3140 CACAGAGATGAGGTGATGCGAGGTGGAATCTTCAACAGATACAAATATCTCAAGATTGAG 3199  
Db 3756 CACAGATGGTGTATGCTGGCGCATCTTCAACAGATACAAATGTCATTGGAATCAA 3815  
QY 3200 AAGTTTGTAAACAGAACTATGGGAAGATACACTCACCGGAGAAAGAAAGTTTCTGAA 3259  
Db 3816 AAGTTTGTAAACAGAACTATGGGAGCGGTCTGCCACCGACAGAGAAAGTCTCTGAG 3875  
QY 3260 GAAACCCACACCATGCAATGAAGAGATGCTATTTATGAGTCTCTCTTTTGTGAATGCA 3319  
Db 3876 GAGAATCAACACCATGCAATGAAGAGATGCTATTTATGAGTCTCTCTTTTGTGAATGCA 3379  
QY 3320 ATTATCCCAAGAGCTTTGATGAAAGGATGCTGATAGTGGTATGTTTGGAGCTGCG 3379  
Db 3936 ATTATTCATAAGGTTTGTATGAGGAGCATGCATACATAGGAGAAATGTTTGGGCGGG 3995  
QY 3380 ATTATTTTGTGAAACTCTTTCCAAAAGCAATCAATATGATATGAAATGGAGGAGGT 3439  
Db 3996 ATTATTTTGTGAAACTCTTTCCAAAAGCAATCAATATGATATGAAATGGAGGAGGT 4055  
QY 3440 ACTGGTGCACAGTTCACAAAGAGATCTGTTTACATTTGCAATTTGCCAGGAGCTCTCTT 3499  
Db 4056 ACAGGCTGCGCTTACACAGAGAGGATGATGCTATATGATGACAGACAAATGCTCTTC 4115  
QY 3500 TGCCGGGTAACTTGGGAAAGCTTTTCTGCAAGTTCAGTGCATGCAATGAAATGGCACTCT 3559  
Db 4116 TGTAGAGTGACCTTGGGAAATCTTTTCTGCAAGTTCAGTGCATGCAATGAAATGGCCACGCG 4175  
QY 3560 CTTCCAGGTCTACCTAGTCACTGCTAGGCGGCTGTAATGGCTAGCATTTAGCTGAA 3619  
Db 4176 CTTCCAGGCGACCTAGTCACTGCTAGGCGGCTGTAATGGCTAGCATTTAGCTGAA 4235  
QY 3620 TATGTTATTTACAGAGGAGCAAGCTTATCTGAGTATTTAATTTACTTACCAGATTATG 3679  
Db 4236 TATGTCATCTACAGAGGAGCAAGCTTATCTGAGTATTTAATTTACTTACCAGATTATG 4295  
QY 3680 AGCCCTGAAG 3689  
Db 4296 AAGCCAGAAG 4305

## RESULT 4

US-09-841-835-9  
; Sequence 9, Application US/09841835  
; Patent No. US2002007695A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4657 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 6...2855  
; US-09-841-835-9

Query Match 34.2%; Score 1297.8; DB 10; Length 4657;  
Best Local Similarity 62.9%; Pred. No. 0;

Matches 2400; Conservative 0; Mismatches 907; Indels 508; Gaps 5;

QY 383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGGGAAACACGTAGTTGAATATTTGCTTCAG 442  
Db 657 TCTCCCCGTCACCTTCGCTGCAGGTTTTGGAAGGAGGATGTTGTAGAACACTTACTACAG 716  
QY 443 AATGGTCAAGTGTCCAAAGCACGATGATGGGGGCTTATCTCTTTCATATGCAATGC 502  
Db 717 ATGGTGTATGCTTCCAGGCTCGTGATGATGGAGTCTCATCCGCTTCATATGCTGT 776  
QY 503 TCTTTTGGTCACTGCTGAAGTAGTCAATCTCTTTTGGACATGCTGCGAGACCCCAATGCT 562  
Db 777 TCTTTTGGCCATGCTGAGTTGTGAGTCTGTTATTTGTCGAAGGAGCTGATCCAAATGCC 836  
QY 563 CGAGATTAATGGAATTTACTCTCTCCATGCAAGCTGCAATTAAGGAAAGATTTGATGTT 622  
Db 837 AGGATAACTGGAATATACACCTCTGCAAGCTGCTATTAAGGAAAGATGATGATGTG 896  
QY 623 TGCATTTGCTGTTACAGCATGGAGCTGAGCAACCATCCGAAATACAGATGGAAGGACA 682  
Db 897 TGCATTTGCTGCTGACGACGAGCTGAGCCCAACACTTCGGAACACTGATGGAAATCA 956  
QY 683 GCATTTGATTTAGCAGATCCATCTGCCAAAGCAGTGTACTTGTGTGAATATAAGAAAGAT 742  
Db 957 GCGCTGGACCTGGCAGATCCTTCAGCAAAAGCTCTCTTACAGGTGAATAAGAAAGAC 1016  
QY 743 GAATCTTTAGAAAGTGCAGGAGTGGCAATGAACAAAAATGATGGCTCTACTACACCA 802  
Db 1017 GAATCTTAGAAGCTGCTAGGAGTGGTAAATGAAGAAAACTTAATGGCTTTTACTGACCT 1076  
QY 803 TTAATGTCAACTGCGAGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTTGGCAGCA 862  
Db 1077 CTAATGTGAATTGCCATGCAAGTGTGGGCGAAAGTCACTCTCTTACATCTAGCAGCG 1136

Qy	863	GGATATTAACAGAGTAAGAATTGTACAGCTGGTTACTTGCACAATGGAGCTGTATGTCCTCATGCT	922
Db	1137	GGCTACACAGAGTGGCAATGGATGTCAGCTTCTTCTTCAGCATGGTGCTGATGTTTCATGCA	1196
Qy	923	AAAGATAAAGGTGATCTGGTACCAATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA	982
Db	1197	AAAGACAAGAGTGAGACTTGTGCCCTTTCATAATGATGTTTCATATGGACAATATTAAGATC	1256
Qy	983	ACTGAACCTTTTGGTCAAGCATGGTGCTGTGTAATGCAATGGCACTTGTGGCAATTCACAT	1042
Db	1257	ACAGAATCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGCTTTACT	1316
Qy	1043	CCTCTTCATGAGGAGCTTCTTAAGACAGGGTTGAAGTATGTTCTTCTTCTCTTAAGTTAT	1102
Db	1317	CCATGCAAGAGCTGCTTCCAAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTATAGCCAT	1376
Qy	1103	GGTGCAGACCCACACATGCTCAATGTGTACAAATAAAAGTGCTATAGACTTGGCTCCACACA	1162
Db	1377	GGCGCTCATCCTACGTTAGTCAACTGCCATGGCAAAAGTGTGTGGATATGGCTCCAACAT	1436
Qy	1163	CCACAGTTAAAGAAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGTGCA	1222
Db	1437	CCGGAGCTTAGGGAGAGATTGACTATGNAITTAAGGTCAATCTTTACTACAAGCAGCC	1496
Qy	1223	CGAGAAGCTGATGTTACTCGAATCAAAAACATCTCTCTGGAATGGTGAATTTCAAG	1282
Db	1497	AGAGAAGCAGACTTAGCTAAAGTTAAAAAACACATCGCTCTGGAATCAITTAATTTCCAA	1556
Qy	1283	CATCCTCAACACATGAACACGATTGCAATGTGCTGTGCATCTCCATGCATCCCAAAAGA	1342
Db	1557	CAACCGAGTCTCATGAACACAGCTGCACATGTGCTGTGGCTCTCTGCATCCCAACAGT	1616
Qy	1343	AAGCAAAATATGTGAAGCTTGTCTAAGAAAAGGAGCAACATCAATGAAAGACATTAAGAA	1402
Db	1617	AACAGTGCAGAAATGTACTTAAAGAAAGGAGCAATGTAATGAAAAAAATAAAGAT	1676
Qy	1403	TTCTTGACTCCTCTGCACGTGGCATCTCAGAAAGCTCAATAATGATGTTGTGAAGTAGTG	1462
Db	1677	TTCATGACTCCCCCTGCATGTCAGCCGAAAGAGCCCAATAATGATGTCATGGAAGTCTG	1736
Qy	1463	GTGAACATCAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA	1522
Db	1737	CATAGCATGGGCCAAGATGAATGCATGGACACCCCTGGTCAGACTGCTTTGTCATAGA	1796
Qy	1523	GCTGCATATTGGTCAATCAAAACCTCGCCGCTACTCCTGAGCTATGGTGTTGATCCT	1582
Db	1797	GCCGCCCTAGCAGGCCACCTGCAGACCTTCGCCCTCTCTGTAGTTAGGCTCTGACCCC	1856
Qy	1583	ACATATTATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAA	1642
Db	1857	TCCATCATCTCCTTACAAGGCTTTCACAGCAGACAGATGGSCAATGAAGCAGTGCAGAG	1916
Qy	1643	CTCCTCCAGAGGGTATCTCATTAGGTAATTTACAGGAGCAGACAGCAATTGCTTGAAGCT	1702
Db	1917	ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTTAGAGGCA	1976
Qy	1703	GCAAGGCTGGAGATGTCGAAACTCTAAAAAACTGTGTAAGTCTGAGAGTGTCACTCC	1762
Db	1977	TCTAAAGCTGGAGACTTGGAACTGTGAAGCACTTTGCAGCTCTCAAAATGGAATTTG	2036
Qy	1763	AGAGACATTTGAAGGGCGTCAGTCTACACCACTTTCATTTGCAAGCTGGTATACAGAGTG	1822
Db	2037	AGAGACTTAGAGGGCGGCAATTCCACGCCCCTTACACTTCGACAGGCTACACCCGCTG	2096
Qy	1823	TCCGTGGTGGATATCTGCTACAGCATGGAGCTGATGTGCATGCTTAAGATAAAGGAGCG	1882
Db	2097	TCGTGTTGAGAGTACCTGCTACACCGGTGCCATGTCATGCCAAAGACAAGGTGCG	2156
Qy	1883	CTTGTAACCTTTGCACAAATGCCATGTTCTTATGGACATTATCAAGTTGCAGAACTCTTGTT	1942
Db	2157	TTGGTGCCCTTCATAATGCCCTGTTTCATATGGACACTTTGAGGTGGCTGAGCTTTTAGTA	2216
Qy	1943	AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAATTTACACCTTTACATGAAGCA	2002

Db	2217	AGGCATGGGGCTTC	TTGTC	CAATG	GGCGGAC	TTAT	GMAATTT	TACCCCT	TC	CAT	GAGCA	2276
Qy	2003	GCAGCAAAAGGAAAT	ATAGAA	TTT	GGCAAACT	TTG	CTCCAGCAT	GGT	GCAGAC	CCCTACC	2062	
Db	2277	GCAGCTAAAGGAAAGT	ATGAA	ATCT	GC	AAGCT	CTTT	AAAA	ACAT	GGAGCAT	CCAACT	2336
Qy	2063	AAAAAAACAGGGAT	GGAAAT	ACT	CCT	TTG	GATCT	TGTT	TAAGAT	GGAGAT	ACAGAT	2122
Db	2337	AAAAAGACAGAGAT	GGAAAT	ACAC	CTT	GGAT	TTGG	TAAAG	GAAGG	AGACAC	AGAT	2396
Qy	2123	CAAGATCTCT	TTAGGGAGAT	GCAGCT	TTGCT	TAGAT	GTGCT	CCAA	GAAGG	TTGTT	TAGC	2182
Db	2397	CAGGACTT	ACTGA	AGGGAT	CTCCT	TTG	TGGAT	GTGCT	CCAA	GAAGG	CTGCT	2456
Qy	2183	AGAGTGAAGAGT	TTCT	CTCT	CGAT	TAAT	GTAA	ATTG	CCCG	CGAT	ACCC	2242
Db	2457	AGAGTGCAGAGCT	CTGT	ACCC	CAGAGA	TAT	CAACT	GC	CAGAG	ACAC	CC	2516
Qy	2243	TCAACACCT	TTACAT	TTAG	CAGCT	GGT	TAT	ATA	ATAT	TTAG	AGTT	2302
Db	2517	TCACCCCT	CTGC	ACCT	GGCAG	CGCT	TAT	ATA	ACCT	GGAA	GTAGCT	2576
Qy	2303	CAACACGGAGCT	GAT	GTGA	ATGCC	CAAGAA	GAGG	AGAC	TAT	TCC	TTT	2362
Db	2577	GAGCATGAGCT	GAT	GTAA	TGCC	CAGACA	AGG	TGGT	TTAA	TCC	TCT	2636
Qy	2363	GCATCTT	ACGG	-----	-----	-----	-----	-----	-----	-----	-----	2373
Db	2637	GCATCTT	ATGGGGCT	GCCT	TGCA	AGAGT	GCAGA	AGCT	CTGT	ACCC	CAGAGAA	2696
Qy	2374	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2373
Db	2697	TGCAGAGAC	CCC	AGGG	CAGAA	TTCA	CCCT	CTGC	ACCT	GGC	AGCAGCT	2756
Qy	2374	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2373
Db	2757	CTGGAAGT	AGCT	GA	TAT	CTT	CTAG	AGCAT	GG	AGCT	GTAT	2816
Qy	2374	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2373
Db	2817	GGTTAA	TTCC	TTT	CAT	AAAT	CGGC	ATCT	TTAT	GGGT	AGT	2876
Qy	2374	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2373
Db	2877	TCCTTT	CAGCT	TTG	TGA	TGAT	TAA	TAG	ACCAT	GC	CAT	2936
Qy	2374	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2373
Db	2937	ATGTAAGCA	TTATA	AAAA	TGCA	GCAT	GTG	TG	CAT	AG	CGCT	2409
Qy	2410	TGCATGT	CTCAAT	GCC	ACG	CAAA	TGGCT	TTT	CAC	ACCT	TTG	2469
Db	2997	CACCTGT	GTAAAT	GCA	ACAGAT	TAAG	TGGG	CGTT	TACT	CCCT	CCAT	3056
Qy	2470	GGGAGCA	ACAGCT	TTT	GTGCT	TTG	TGCT	TAG	CCCAT	GGAG	CTG	2529
Db	3057	AGGAAG	CAGCAGCT	GTG	CGCCCT	CTCCT	CTAG	CGCAT	GGT	GCAG	ACCC	3116
Qy	2530	TCAGGA	AGGACAA	AC	CTTT	TAGAT	TTAG	TTT	CAG	CGGAT	GTAT	2585
Db	3117	CCAGGA	AGGCC	CAG	CGCT	CTGG	ATCT	GGC	ACAGCT	GCAG	TAT	3176
Qy	2586	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2585
Db	3177	TGACAT	AGCGCT	TTT	TAT	TGATA	AAAA	TACA	CAC	GTG	TGTA	3236
Qy	2586	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2585
Db	3237	GTTT	TACT	CC	CCCT	CCAT	GAA	GCAG	CCCA	GCAG	CTGT	3296
Qy	2612	CTGCC	CTCT	TTGTT	ACA	GCCT	CAAG	-----	-----	-----	-----	2636



Db 3297 AGCGCATGTGCGAGACCCACCACCAAGAAACAGGAAGCCAGAGCGCTCTGGATCTGGC 3356  
QY 2637 -----TGCTCAATGTTGAGAAACCCAGGAGCCACTGC--- 2670  
Db 3357 AACAGCTGACGATATCAGAGCTTTGCTGATAGCCATGCCCCAGAGGCGCTTACCCTAC 3416  
QY 2671 -----AGATGCTCTCTCTTCCAGGTCCATCTAGGCC 2700  
Db 3417 CTGCTTTAAACCTCAGGCTACTGTAGTAGTGCCTCTCTGATCTCACCAGCATCCACCCC 3476  
QY 2701 ATCAAGCCTTTTCGACGACGAGCTTTGACAACATATCTGGAGTFTTTTCAGAACTGTC 2760  
Db 3477 CCTCTGCTCTGGCTGCGGTCGACGAGCATAGACAACCTCACTGGCCCTTTAGCAGAGTTGC 3536  
QY 2761 TTCAGTAGTGTAGTTCAGTGGACAGAGGTCTCCAGTTTGGGAGAAAAG-----CA 2814  
Db 3537 CGTAGGAGAGCCCTCAATGCGAGGGATGGCGCGGGAACAGAAAGGAAGGAGA 3596  
QY 2815 GPTTCCAGGAGTAGATTTTAGCATACTCAATTCGTAAAGAACTCTTGGACTTGGACACT 2874  
Db 3597 AGTTGCTGCTTGGACATGAATATCAGCCAATTTCTTAAAAAGCCTTGGCCCTTGAACACT 3656  
QY 2875 AATGGATATATTGACAGAGACAGATCACTTTGGATGTATTAGTTGAGATGGGGACAA 2934  
Db 3657 TCGGGATATCTTTGAAACAGACAGATTACACTAGATGTGTTGGCTGATATGGGTCACTGA 3716  
QY 2935 GGAGCTGAAGGAGATTGGAACTCAATCTTATGGACATAGCACAAACTAAATTAAGGAGT 2994  
Db 3717 AGAGTTGAAGAATAAGGCATCAATGCATATGGCCACCCCAAAATTAATCAAGGAGT 3776  
QY 2995 CGAGAGACTTATCTCGGACCAAGGCTTAAACCCATATTTAACTTTGAACACCTCTGG 3054  
Db 3777 AGAAGACTCTTAGTGGACAAAGGACCAACCTTATTTCACTTTTCACTGTCTTAA 3836  
QY 3055 TAGTGGAACTTCTTATAGATCTGCTCTGATGATATAAGAGTTTCACTCTGTGGAGGA 3114  
Db 3837 TCAGGGAAGCATTTTCTGCTGATCTTCTCCAGAGAATAAAGAAATATCAGTCAGTGAAGA 3896  
QY 3115 AGAGATGCAAGTACAGTTTCGAGAGCACAGATGAGGTCTGAGGTGGAATCTTCAA 3174  
Db 3897 AGAGATGCAAGTACATTCGAGAAACAGAGATGTTGTAATGCTGGCGCATCTTCAA 3956  
QY 3175 CAGATCAATATTTCTCAAGATTCAGAGGTTTGTAAACAAAGAACTATGGGAAAGATACAC 3234  
Db 3957 CAGATCAATATGTCGAATTCAAAAGTTGTCAACAAGAGTTGAGGGAGCGGTTCTG 4016  
QY 3235 TCACCGGAGAAAAGAGTTTTCGAAAGAACACACACATGCCATGCCAATGACAGATGCTATT 3294  
Db 4017 CCACCGACAGAGGAAGTGTCTGAGGAGAAATCACACCATCAATGAGCGCATGTTGT 4076  
QY 3295 TCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTGATGAAGGATCGTA 3354  
Db 4077 TCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTGATGAAGGATCGTA 4136  
QY 3355 CATAGTGTGATGTTTGGAGCTGGCATTTTATTTGCTGAAAACCTCTTCCAAAAGCAATCA 3414  
Db 4137 CATAGGAGGAATCTTTGGGCGCGGATTTATTTTGTGAAAACCTCTCTCAAAAAGCAACCA 4196  
QY 3415 ATATGATATGGAATTTGGAGGAGTACTGGGTGTCAGATTCACAAAGACAGATCTTGTTA 3474  
Db 4197 ATATGTTTATGGAATTTGGAGGAGGAACAGGCTGCCCTTACACACAAAGGATCTATGCTA 4256  
QY 3475 CATTTGGCCACAGCAGCTGCTCTTTTGGCGGTTAACTTTGGGAAAGCTTTCTTCGCGATT 3534  
Db 4257 TATATGTCACAGAAATGCTCTCTCTGTAGAGTGACCCCTTGGGAAATCTCTTCTGCGATT 4316  
QY 3535 CAGTGAATGAAATGGCACAATTTCTCTCAGGTCTATCACTCAGTCACTGAGGCGCCAG 3594  
Db 4317 TAGCACCATGAAATGGCCACCGCTCCAGGGCACCCTCAGTCACTTGGTAGACCGAG 4376  
QY 3595 TGTAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGCTTATCCTGA 3654  
Db 4377 CGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATATCCAGA 4436

QY 3655 GTATTTTATTTACCATGATTATGAGCCCTGAAG 3689  
Db 4437 GTATCTTATCTTACCATGATTATGAGCCAGAAG 4471

## RESULT 5

US-09-964-899-40  
; Sequence 40, Application US/09964899  
; Patent No. US2002017446A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Dalia et al.  
; TITLE OF INVENTION: Identification of Genes Involved in  
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster  
; FILE REFERENCE: 4-31612 A  
; CURRENT APPLICATION NUMBER: US/09/964,899  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/236,893  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/298,309  
; PRIOR FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 2409  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-964-899-40

Query Match 26.2%; Score 994; DB 9; Length 2409;  
Best Local Similarity 65.4%; Pred. No. 1.8e-254;  
Matches 2174; Conservative 0; Mismatches 20; Indels 1128; Gaps 4;

QY 403 AGCTTTTGGGCGGAAGAGCTAGTTGAATATTTGCTTCAGAAATGGTCCAAGTGTCCAAAGC 462  
Db 198 AGGTTTGGGCGGAAGAGCTAGTTGAATATTTGCTTCAGAAATGGTCCAAGTGTCCAAAGC 257  
QY 463 AGCTGATCATGGGGCCCTTATTCCTCTTCAATAATGATGCTCTTTTGGTCAATGCTGAAGT 522  
Db 258 AGCTGATCATGGGGCCCTTATTCCTCTTCAATAATGATGCTCTTTTGGTCAATGCTGAAGT 317  
QY 523 AGTCAATCTCCTTTTGGACATGTTGACAGCCCAATGCTCGAGATAATTTGGAATATATAC 582  
Db 318 AGTCAATCTCCTTTTGGACATGTTGACAGCCCAATGCTCGAGATAATTTGGAATATATAC 377  
QY 583 TCCTCTCCATGAAGTGAATTAAGGAAAGATTTGATGCTTTGCTGCTGTACAGCA 642  
Db 378 TCCTCTCCATGAAGTGAATTAAGGAAAGATTTGATGCTTTGCTGCTGTACAGCA 437  
QY 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 702  
Db 438 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATCC 497  
QY 703 ATCTGCCAAAGCAGTGTCTTACTGTGGAATAT-----AAGAAGATGA 744  
Db 498 ATCTGCCAAAGCAGTGTCTTACTGTGTAAGTCTGTATACCTCTGTATCTTATCCAGGAAGCCTGT 557  
QY 745 ACTCTTAGAAAGTGCAGAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTT 804  
Db 558 AAAGAACACCTTGGCAGGAGTGGCAATGAAGAAAATGATGGCTCTACTCACACCATTT 617  
QY 805 AAATGCTCAACTGCCACGAAGTGTATGGCAGAGAAAGTCAACTCCATTTACATTTTGGCAGCAGG 864  
Db 618 AAATGCTCAACTGCCACGAAGTGTATGGCAGAGTGGCAGAG----- 648  
QY 865 ATATAACAGAGTAAAGATTTGTACAGCTGTTTACTGCCAATGGAGCTGATGCTCATGCTAA 924  
Db 649 ----- 648  
QY 925 AGATAAAGGTGATCTGGTACCATTACACAATGCCCTGTTTATGCTCAATTTAAGATTAAC 984  
Db 649 ----- 648









US-09-964-899-42  
; Sequence 42, Application US/09964899  
; Patent No. US2002017446A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Dalia et al.  
; TITLE OF INVENTION: Identification of Genes Involved in  
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster  
; FILE REFERENCE: 4-31612 A  
; CURRENT APPLICATION NUMBER: US/09/964,899  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/236,893  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/298,309  
; PRIOR FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42  
; LENGTH: 5175  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-964-899-42

Query Match 2.1%; Score 80.8; DB 9; Length 5175;  
Best Local Similarity 45.5%; Pred. No. 5.4e-11;  
Matches 390; Conservative 0; Mismatches 447; Indels 21; Gaps 2;

QY 1356 AACTGTTCTGAAGAAAGAGCAACATCAATGAAAGACTAAAGAAATTCCTGACTCCTC 1415  
Db 740 AACTATTCTCGATCGAGGAGTAAATCGATGCCAAACACGAGGATGGTCTGACACAC 799  
QY 1416 TGCAGTGGCATCTGAGAAGCTCATATGATGTGTGTTGAAGTAGTGTGTAACATGAAG 1475  
Db 800 TGCAGTGGAGCAAGGAGTGGCCACGACGAGGTGGTGAAGATGTTGTTGATCGAGCTG 859  
QY 1476 CAAAGGTTATGCTCTGATAATCTTGCTCAGACTTCTTACACAGACTGCATATTG 1535  
Db 860 CCCCATTCTTCAAACCAAGATGATTTATCTCCATTGCATGCCACACAAAGGG 919  
QY 1536 GTCACTACAACCTGGCCCTACTCTGAGCTATGGGTGTGATGCTTAACATTATATCC 1595  
Db 920 ATCATTTAACTGGCTCCAGCTTCTCTCCAGCATATGTACCCGTGGATGATGTCACCA 979  
QY 1596 TTCAGGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCACTCTCCAAGAG 1655  
Db 980 ATGACTACCTGACCTGCTTACAGCTGGCTGCCACTGTGGCCATTACAAAGTTCGCAAG 1039  
QY 1656 GTATCTCATTAAGTAATTCAGAGGCAGACACAAATTCCTGGAAGCTGCAAGGCTGGAG 1715  
Db 1040 TTCTCTTGGATAAGAAAGCTAACCCTAATGCCAAGCCCTGAATGGCTTTTACCCTCTTC 1099  
QY 1716 ATGTCGAA-----ACTGTAAAAAAGCTGTGACTGTGTTTCAGAGTGTCA 1757  
Db 1100 ATATTGGCTGCAAGAAATCGAATTAAGTAATGGAATCTCTTCTGAAACACGCTGCAT 1159  
QY 1758 ACTCAGAGACATTTGAAGGGCTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1817  
Db 1160 CCATCCAAAGCTGTAAACGAGAGAGAAACAGACATGCACATGGCAGCTCGCTCCGGCC 1219  
QY 1818 GAGTGTCCGGTGGATATCTGCTACAGCATGAGCTGATGTGCATGCTTAAGATAAAG 1877  
Db 1220 AAGCTGAAGTGTGCGGTATCTGTTACAAAGAGGAGCTCAGGTAGAAAGCTAAAGCTAAG 1279  
QY 1878 GAGGCTTTGACCTTTGCACAATCATGTTCTTATGACATATTAAGTGTGCAAACTTC 1937  
Db 1280 ATGACCAACACCACTCCCAATTCAGCCCGACTGGGGGAAGCAGACATAGTACAAAGC 1339  
QY 1938 TTGTTAAACATGGAGCAGTAGTAAATAGTGTATTTATGAAATTTTACACCTTTACATG 1997  
Db 1340 TGTTCAGCAAGGGCATCTCCAAATCAGCCACAACTTCTGGGTACACCCCACTTCACC 1399  
QY 1998 AAGCAGCAGCAAAAGGAAATATGAATTTGCAAACTTCTGCTCCAGCAGTGGTGCAGACC 2057  
Db 1400 TTTCCCGCCGAGGGGCGATGAGGATGGCCGGTTCCTTTTGGATCATGGAGCGTCTT 1459

QY 2058 CTACCAAAAAAACAGGGATGGAATACTACTCCTTTGATCTTGTAAAGATGGAGATACAG 2117  
Db 1460 TATCTATATACAAAGAGCGGGCTAACACCACTGCATGTAGCTGCACATTTACGATAATC 1519  
QY 2118 ATA---TTCAAGATCTGTTAGGGAGATGACAGCTTTTGTAGATGCTGCCAAGAGGGTT 2174  
Db 1520 AGAAAGTGCCTTCTGCTTTTGGACCAAGAGGAGCTCAGCTCAGCGCGCGCAAGAATG 1579  
QY 2175 GTTTAGCCAGAGTGAAGA 2192  
Db 1580 GTTATACGCCACTGCACA 1597

RESULT 13  
US-09-908-711-12  
; Sequence 12, Application US/09908711  
; Patent No. US20020045230A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P128  
; CURRENT APPLICATION NUMBER: US/09/908,711  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US01/01360  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,867  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01344  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,892  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01345  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,888  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01329  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,905  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01354  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,891  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01339  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,869  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01340  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,874  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01334  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,898  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01320  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,853  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01349  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,902  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01239  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,870  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01348  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 09/764,882  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01347





Db 586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATT 645  
QY 2027 TGCAAACTTCTGCTCCAGCATGG 2049  
Db 646 GCAAACTCTTGTGGAAGAAGG 668

## RESULT 15

US-09-947-199-1  
; Sequence 1, Application US/09947199  
; Patent No. US20020127684A1  
; GENERAL INFORMATION:  
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: MNI-068CP2  
; CURRENT APPLICATION NUMBER: US/09/947,199  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: 60/111,938  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/291,839  
; PRIOR FILING DATE: 1999-04-14  
; PRIOR APPLICATION NUMBER: 09/458,457  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3025  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (48)..(2552)  
US-09-947-199-1

Query Match 2.08; Score 74.2; DB 10; Length 3025;  
Best Local Similarity 55.18; Pred. No. 2.3e-09;  
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
QY 1787 ACACCACCTTCATTTTGCAGCTGGGTATACACAGAGTGCCGGTGGGAATATCTGCTACAG 1846  
Db 453 ACTGCCCTCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGTGCTGTGTGCAA 512  
QY 1847 CATGGAGCTGATGCTGATGCTAAAGATAAAGAGGCCCTTGTACCTTTGCACAATGCATGT 1906  
Db 513 CATGGAGCTAATGCAATATTCAGATGCAGTCTTTTCACTCCCATTCATATTCACGG 572  
QY 1907 TCTTATGACATTTATGAAGTTGCACACTTCTTGTAAACATGGAGCAGTAGTAAATGTA 1966  
Db 573 TACTATGGACATGAACAGGTAACCTGCGCTCTTTTGAATTTGGTCTGTATGTAATGTA 632  
QY 1967 GCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATT 2026  
Db 633 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATT 692  
QY 2027 TGCAAACTTCTGCTCCAGCATGG 2049  
Db 693 GCNAAACTCTTGTGGAAGAAGG 715

Search completed: February 12, 2003, 03:31:24  
Job time : 213.581 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 12:18:03 ; Search time 4936.65 Seconds  
(without alignments)  
12456.682 Million cell updates/sec

Title: US-09-843-159B-1

Perfect score: 3797

Sequence: 1 ctttgagacactggatttc.....ccttgctgaaaaa 3797

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	784.2	20.7	838 9 AU140145
2	672.2	17.7	1128 13 BM457025
3	649.2	17.1	659 14 R64714
4	643.8	17.0	691 12 BG391376
5	642.8	16.9	759 9 AL601027
6	634.4	16.7	950 14 BQ885764

7	605	15.9	648	12	BG875327	RC1-BT062
8	593.2	15.6	622	10	BE220005	h65h03.x
9	584	15.4	616	10	BE467629	h65b11.x
10	570.4	15.0	956	12	BE910901	601661978
11	568	15.0	702	10	BB339554	BB339554
12	562.8	14.8	585	10	BE222311	hul0d01.x
13	557.4	14.7	595	10	BE504175	h586c11.x
14	549	14.5	553	14	BQ267594	ij90c10.y
15	544	14.3	1099	12	BG245969	602358786
16	538.6	14.2	562	10	BE504478	h257a02.x
17	534.4	14.1	956	12	BE585334	602102008
18	518.4	13.7	520	10	AW157349	au94h04.x
19	504	13.3	729	13	BG917245	602816521
20	503.4	13.3	553	12	BE721392	188553 MA
21	487.4	12.8	501	14	BM751146	K-EST0027
22	474.4	12.5	984	12	BF384484	602046448
23	462.2	12.2	600	12	BG807796	2072-17 M
24	459	12.1	460	9	AA307492	EST178376
25	456	12.0	748	14	BQ179839	UI-M-BW0-
26	450.8	11.9	645	10	BE373823	601227836
27	446.8	11.8	597	10	AW749617	RC3-BT050
28	437.4	11.5	1094	13	BM458871	AGENCOURT
29	431.8	11.4	441	10	AW208871	UI-H-B11-
30	430.6	11.3	580	13	BI064735	pgfln.pk0
31	420.4	11.1	709	13	BI152307	BI152307
32	413.4	10.9	497	10	AW231078	uo70e07.y
33	405.8	10.7	469	10	BE503173	h284e01.x
34	401.8	10.6	598	9	AL680319	AL680319
35	399	10.5	411	9	AL711456	DKF2p6860
36	388	10.2	412	9	AL711468	DKF2p686P
37	380.4	10.0	384	9	AA244138	nc04g08.s
38	374	10.0	391	9	AA244137	nc04g08.r
39	377.6	9.9	458	12	BF021623	uy49g12.y
40	374.8	9.9	677	13	BJ038772	BJ038772
41	366.2	9.6	878	12	BF979977	602288547
42	362.8	9.6	819	13	BI558956	60324317
43	362	9.5	660	10	BB391900	BB391900
44	360	9.5	1114	13	BM464678	AGENCOURT
45	354.6	9.3	564	14	HI7748	ym40g08.r1

# ALIGNMENTS

RESULT 1	AU140145	838 bp	mRNA	linear	EST 05-AUG-2002
LOCUS	AU140145	PLACE2	Homo sapiens	cdna	clone PLACE20000021 5', mRNA
DEFINITION	sequence.				
ACCESSION	AU140145				
VERSION	AU140145.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y. and Isogai, T.				
TITLE	HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y., Isogai, T.)				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and				

FEATURES	Helix Research Institute.		Location/Qualifiers		RESULT 2	
	source	1. 838	/organism="Homo sapiens"	/db_xref="taxon:9606"	BM457025	LOCUS
		/clone="PLACE2000021"		AGENCOURT_6406851 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583257		DEFINITION
		/clone_lib="PLACE2"		5', mRNA sequence.		ACCESSION
		/tissue_type="placenta"		BM457025		VERSION
		/note="Vector: pME18SFL3"		BM457025.1 GI:18506065		KEYWORDS
				EST.		SOURCE
				human.		ORGANISM
				Homo sapiens		
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
				Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
				NIH-MGC http://mgc.nci.nih.gov/.		
				National Institutes of Health, Mammalian Gene Collection (MGC)		
				Unpublished (1999)		
				Contact: Robert Strausberg, Ph.D.		
				Email: cgapbs-r@mail.nih.gov		
				Tissue procurement: ATCC		
				cDNA Library Preparation: Life Technologies, Inc.		
				cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
				DNA Sequencing by: Agencourt Bioscience Corporation		
				Clone distribution: MGC clone distribution information can be		
				found through the I.M.A.G.E. Consortium/LLNL at:		
				http://image.llnl.gov		
				Plate: LLAM12346 row: a column: 18		
				High quality sequence stop: 507.		
				Location/Qualifiers		
				1. .1128		
				/organism="Homo sapiens"		
				/db_xref="taxon:9606"		
				/clone="IMAGE:5583257"		
				/clone_lib="NIH_MGC_92"		
				/tissue_type="embryonal carcinoma, cell line"		
				/lab_host="DH10B (phage-resistant)"		
				/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;		
				Site_2: SalI; Cloned unidirectionally; oligo-dT primed.		
				Average insert size 2.5 Kb. Library enriched for		
				full-length clones and constructed by Life Technologies.		
				Note: this is a NIH_MGC Library."		
				328 a 297 c 232 g 271 t		
				BASE COUNT		
				ORIGIN		
				Query Match 17.7%; Score 672.2; DB 13; Length 1128;		
				Best Local Similarity 97.7%; Pred. No. 3.1e-153;		
				Matches 724; Conservative 0; Mismatches 13; Indels 4; Gaps 4;		
				QY 1957 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAGGAA 2016		
				Db 1 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAGGAA 60		
				QY 2017 ATATGAAATTTGCAAACTTCTGCTCCAGCATGTCGAGACCTTACCACCAAAAAACAGGGA 2076		
				Db 61 ATATGAAATTTGCAAACTTCTGCTCCAGCATGTCGAGACCTTACCACCAAAAAACAGGGA 120		
				QY 2077 TGGAAATACCTCTTGGATCTTGTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 2136		
				Db 121 TGGAAATACCTCTTGGATCTTGTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 180		
				QY 2137 GGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 2196		
				Db 181 GGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 240		
				QY 2197 GTCCTCTCTGATATAATGTAATTCGCCGATACCCAGCAGACATTCACACCTTTTACA 2256		
				Db 241 GTCCTCTCTGATATAATGTAATTCGCCGATACCCAGCAGACATTCACACCTTTTACA 300		
				QY 2257 TTTCAGAGCTGGTTATAATTAATTTAGAAAGTTCAGAGTATTGTTTACAACAGGAGCTGA 2316		
				Db 301 TTTCAGAGCTGGTTATAATTAATTTAGAAAGTTCAGAGTATTGTTTACAACAGGAGCTGA 360		
				QY 2317 TGTGAATGCCCAGACAAAGGAGGACTTATCTCTTACATATATGACAGATCTTACGGGCA 2376		

```

Db 361 TGTGAATGCCCAAGACAAAGAGGAGCTATTTCCTTTACATAATGACGACATCTTACGGGCA 420
QY 2377 TGTAGATGTAGCAGCTCTACTATAAAGTATAATGATGTGTCAATGCCCGGACAAATG 2436
Db 421 TGTAGATGTAGCAGCTCTACTATAAAGTATAATGATGTGTCAATGCCCGGACAAATG 480
QY 2437 GGCCTTTCACACCTTTGCGAGGAGAGCCCAAAAGGAGCAGACACAGCTTTGTGCTTTGTT 2496
Db 481 GGCCTTTCACACCTTTGCGAGGAGAGCCCAAAAGGAGCAGACACAGCTTTGTGCTTTGTT 540
QY 2497 GCTAGCCATGAGCTGAGCCGACCTTTAAATAATCAGGAAGGACAAACA-CCTTTAGATT 2555
Db 541 GCTAGCCATGAGCTGAGCCGACCTTTAAATAATCAGGAAGGACAAACACCCCTTTACATT 600
QY 2556 TAGTTTTCAGCGGATGATGTCAGCG-CTCTTCTGACAGCAGCATGCCCCCATCTGCTCTG 2614
Db 601 TAGTTTTCAGCGGATGATGTCAGCGCTCTTCTGACACCGCATGGCCCATCTGCTCTG 660
QY 2615 CCCTCTTTTACAAGCCTCAAGTGTCTCAATGGTG-TGAGAGCCCGAGGAGCCACTGCGAGA 2673
Db 661 CCCTCTTTGTCACCGCTCAAGCGCTCCATGGTGTGAGAGCCACGCGCCACTGCGAGA 720
QY 2674 TG-CTCTCTCTTCAGTGCCAT 2693
Db 721 TGCTTTCTCTTCAGGTTCTCT 741

RESULT 3
R64714
LOCUS
DEFINITION
EST54409 WATM1 Homo sapiens cDNA clone 54a09 similar to BLASTX
homology with sp|Q01485|ANK_HUMAN ANKYRIN, BRAIN VARIANT 2, mRNA
sequence.
R64714
VERSION
R64714.1 GI:837352
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 659)
Bouillaud,F.
Study of expressed sequences tags in adipose tissue 1995
Unpublished (1995)
Contact: Frederic Bouillaud
Centre de Recherche sur l'Endocrinologie moleculaire et le
Developpement
CNRS
9, Rue Jules Hetzel, Meudon Bellevue, 92190 France
Tel: 33 1 45 07 52 87
Fax: 33 1 45 07 58 90
Email: bouillaud@infobiogen.fr
automatic cycle sequencing of PCR amplified insert, open reading
frame 3. 659.
FEATURES
Location/Qualifiers
source
1. 659
/organism="Homo sapiens"
/strain="Caucasian"
/db_xref="taxon:9606"
/clone="54a09"
/clone_lib="WATM1"
/note="Vector: lambda gtl; Site_1: EcoRI; White adipose
tissue, sub cutaneous, adult, female. Purification of
polya mRNA, first strand priming with random
oligonucleotides. Ligation of adaptors EcoRI/NotI, cloning
in EcoRI site of lambda gtl."
200 a 128 c 154 g 172 t 5 others
BASE COUNT
ORIGIN
Query Match 17.1%; Score 649.2; DB 14; Length 659;
Best Local Similarity 98.6%; Pred. No. 1.1e-147;
Matches 649; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1416 TGCACGTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGCTGAAACATGAAG 1475
Db 1 TGCACGTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGCTGAAACATGAAG 60
QY 1476 CAAAGGTTAATGCTCTGGATAAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTGTG 1535
Db 61 CAAAGGTTAATGCTCTGGATAAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTGTG 120
QY 1536 GTCATCTACAAACCTGCCGCTACTCCTGAGCTATGGTGTGATCCCTAACATTATATCCC 1595
Db 121 GTCATCTACAAACCTGCCGCTACTCCTGAGCTATGGTGTGATCCCTAACATTATATCCC 180
QY 1596 TTCAGGCTTTTACTGCTTTACAGATGGAAATGAAAATGTACAGCAACTCCTCCAAGAGG 1655
Db 181 TTCAGGCTTTTACTGCTTTACAGATGGAAATGAAAATGTACAGCAACTCCTCCAAGAGG 240
QY 1656 GTATCTCATTTAGTAAATTCAGAGGCACAGACAAATTTGCTGGAGCTGCAAGGCTGGAG 1715
Db 241 GTATCTCATTTAGTAAATTCAGAGGCACAGACAAATTTGCTGGAGCTGCAAGGCTGGAG 300
QY 1716 ATGTCGAAACTGTAAAAAACTGTGTACTGTTCAAGCTGTCAACTGCAGAGACATTGAAG 1775
Db 301 ATGTCGAAACTGTAAAAAACTGTGTACTGTTCAAGCTGTCAACTGCAGAGACATTGAAG 360
QY 1776 GGCCTCAGTGTACACCACTTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAAT 1835
Db 361 GGCCTCAGTGTACACCACTTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAAT 420
QY 1836 ATCTGCTACAGATGGAGCTGATGTGATGCTAAAGATAAAGAGGCTTGTACCTTTGC 1895
Db 421 ATCTGCTACAGATGGAGCTGATGTGATGCTAAAGATAAAGAGGCTTGTACCTTTGC 480
QY 1896 ACAATGATGTTCTTATGAGACATTATCAAGTGCAGAACTTCTTGTAAACATGGAGCAG 1955
Db 481 ACAATGATGTTCTTATGAGACATTATCAAGTGCAGAACTTCTTGTAAACATGGAGCAG 540
QY 1956 TAGTTAATGTAGTGTATTTATGAAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAA 2015
Db 541 TAGTTAATGTAGTGTATTTATGAAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAA 600
QY 2016 AATATGAAATTTGCAAACTTCTGCTCCAGCATGTGTCAGACCCCTACCAAAAAAACAG 2073
Db 601 AATATGAAATTTGCAAACTTCTGCTCCAGCATGTGTCAGACCCCTACCAAAAAAACAG 658

RESULT 4
R6391376
LOCUS
DEFINITION
602417447F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4537138 5',
mRNA sequence.
R6391376
ACCESSION
R6391376.1 GI:13284824
VERSION
R6391376.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 691)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMI0461 row: m column: 11
High quality sequence stop: 684.
Location/Qualifiers
FEATURES

```

source	1..691					ACCESSION	AL601027
	/organism="Homo sapiens"					VERSION	AL601027.1
	/db_xref="taxon:9606"					KEYWORDS	GT:15164533
	/clone="IMAGE:4537138"					EST.	human.
	/clone_lib="NIH_MGC_92"					SOURCE	Homo sapiens
	/tissue_type="embryonal carcinoma, cell line"					ORGANISM	
	/lab_host="DH10B (phage-resistant)"					REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;					AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
	Site_2: SalI; Cloned unidirectionally; oligo-dT primed.						1 (bases 1 to 759)
	Average insert size 2.5 kb. Library enriched for						Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
	full-length clones and constructed by Life Technologies.						.S.
	Note: this is a NIH_MGC Library."					TITLE	EST (Duesterhoeft, et al.)
BASE COUNT	218 a 122 c 165 g 186 t					JOURNAL	Unpublished (1999)
ORIGIN						COMMENT	Contact: Duesterhoeft A
	Query Match 17.0%; Score 643.8; DB 12; Length 691;						MIPS
	Best Local Similarity 99.3%; Pred. No. 2.3e-146;						Am Klopferspitz 18a D-82152 Martinsried, Germany
	Matches 678; Conservative 0; Mismatches 2; Indels 3; Gaps 3;						This is the 5' sequence of the clone insert
QY	1706 AAGGCTGGAGATGTCGAAACTGTAAAAAAGTGTACTGTTTCAGAGTGTCAACTGCAGA 1765						Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Db	10 AAGGCTGGAGATGTCGAAACTGTAAAAAAGTGTACTGTTTCAGAGTGTCAACTGCAGA 69						Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
QY	1766 GACATTGAAGGGCTCAGTCTACACACATTCATTTTGCAGCTGGGTATACAGAGTGTCC 1825						sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
Db	70 GACATTGAAGGGCTCAGTCTACACACATTCATTTTGCAGCTGGGTATACAGAGTGTCC 129						consortium of the German Genome Project.
QY	1826 GTGGTGGAAATATCTGTACAGATGAGCTGATGTCATGCTAAAGATAAAGGAGGCCCTT 1885						No sl sequence available.
Db	130 GTGGTGGAAATATCTGTACAGATGAGCTGATGTCATGCTAAAGATAAAGGAGGCCCTT 189						This clone (DKFZp313G2239) is available at the RZPD in Berlin.
QY	1886 GTACCTTTGCACAATGCATGTTCTTATGACATATATGAAGTTCGACAATCTTGTATAA 1945						Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Db	190 GTACCTTTGCACAATGCATGTTCTTATGACATATATGAAGTTCGACAATCTTGTATAA 249						Berlin- Charlottenburg, GERMANY; Email: cloneerzpd.de.
QY	1946 CATGGAGCAGTAGTAACTAGCTGATTTATGGAATTTACACCTTTACATGAGCAGCA 2005					FEATURES	Location/Qualifiers
Db	250 CATGGAGCAGTAGTAACTAGCTGATTTATGGAATTTACACCTTTACATGAGCAGCA 309					source	1..759
QY	2006 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCA-TGGTGCAGACCCCTACCAA 2064						/organism="Homo sapiens"
Db	310 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCACTGGTGCAGACCCCTACAA 369						/db_xref="taxon:9606"
QY	2065 AAAAAACAGGATGGAATACTCCTTTGGATCTGTTAAAGATGGAGATACAGATATTC 2124						/clone="DKFZp313G2239"
Db	370 AAAAAACAGGATGGAATACTCCTTTGGATCTGTTAAAGATGGAGATACAGATATTC 429						/clone_lib="313 (synonym: hlcc2)"
QY	2125 AGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTAGCCAG 2184						/dev_stage="adult"
Db	430 AGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTAGCCAG 489						/lab_host="DH10B"
QY	2185 AGTGAAGAAGTTGCTCTCCCTGATA-ATGTAAATTCGCGCGATACCCCAAGGCAGACATT 2243						/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
Db	490 AGTGAAGAAGTTGCTCTCCCTGATACATGTAAATTCGCGCGATACCCCAAGGCAGACATT 549						cDNA-collection"
QY	2244 CAACACCTTTACATTTAGCAGCTGGTTAATAAATTTAGAAGTTGCAGAGTATTTGTTAC 2303					BASE COUNT	217 a 156 c 175 g 201 t 10 others
Db	550 CAACACCTTTACATTTAGCAGCTGGTTAATAAATTTAGAAGTTGCAGAGTATTTGTTAC 608					ORIGIN	
QY	2304 AACACGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 2363						Query Match 16.9%; Score 642.8; DB 9; Length 759;
Db	609 AACACGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 668						Best Local Similarity 97.2%; Pred. No. 4.2e-146;
QY	2364 CATCTTTACGGGCATGTAGATGTA 2386						Matches 669; Conservative 0; Mismatches 17; Indels 2; Gaps 2;
Db	669 CATCTTTACGGGCATGTAGATGTA 691						
RESULT 5							
LOCUS	AL601027						
DEFINITION	DKFZp313G2239_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone						
	DKFZp313G2239_5', mRNA sequence.						

```
QY 2854 GAATCTTGAGCTTGAGCACCTAAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 2913
Db 552 GAATCTTGAGCTTGAGCACCTAAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 611
QY 2914 ATTACTTCAGATGGGCGACAGGAGCTGAAGAGATTTGGAATCAATCTTATGACATAG 2973
Db 612 ANAGTTGAGATGGGCGACAGGAGCTGAAGAGATTTGGAATCAATCTTATGACATAG 671
QY 2974 GCACAACTAATTTAA-GGAGTCGAGAGACTTATCTCCGGACAACAA-GGTCTTTAAACCCA 3031
Db 672 GCACAACTAATTTAAAGGAGTCGAGAGACTNATCTCCGGACAACAGGGGCTTTAAACCCA 731
QY 3032 TATTTAACTTTGAACACCTCTGGTAGTG 3059
Db 732 TATTTAACTTTGAACACCTCTGGNAGAG 759

RESULT 6
BQ885764
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BQ885764 950 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8746924 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6392888
5', mRNA sequence.
BQ885764
BQ885764.1 GI:22277782
EST.
house mouse.
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 950)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM13883 row: h column: 09
High quality sequence stop: 641.
FEATURES
source
1..950
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6392888"
/clone_lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: otcysts; Vector: pCMV-SPORT6.1.cdb;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dT. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
BASE COUNT 280 a 195 c 226 g 245 t 4 others
ORIGIN

Query Match 16.7%; Score 634.4; DB 14; Length 950;
Best Local Similarity 86.4%; Pred. No. 5.2e-144;
Matches 759; Conservative 0; Mismatches 111; Indels 8; Gaps 5;

QY 2481 AGCTTTGCTTTGTTGCTAGCCCATGAGCTGACCCGACTTTAAATCAGGAAGGAC 2540
Db 1 AGCTTTGCTTTGTTGTTGCTAGCCCATGAGCTGACCCGACTTTAAATCAGGAAGGAC 59
QY 2541 AAACACCTTTAGATTTTACGGGATGATGTCAGGCGCTTCTGACAGCAGCCATGC 2600
Db 60 AAACACCTTTAGATTTTACGGGATGATGTCAGGCGCTTCTGACAGCAGCCATGC 119
QY 2601 CCCCATCTGCTGCCCCCTTTGTTTACAAGCCTCAAGTCTCAATGGTGTGAGAGCCGAG 2660
```

```
Db 120 CCCCCCTGCTGCTGCTACGTGCTACAAACCTCAAGTGCCTCAGTGGCGTGAGGGGCCCCG 179
QY 2661 GAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTCTCTGCAGCCA 2720
Db 180 GAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTCTCTGCAGCCA 239
QY 2721 GCAGTCTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTTCAGTAGTGTAGTTCAGTG 2780
Db 240 GCAGCCTCGAACACTTATCTGGCAGCTTCTCGGAACCTGTCGCGAGTGGTGTAGTTCAGTG 299
QY 2781 GAACAGAGGTGCTTCCAGTTTGGAGAAAGAGGTTCCAGGAGTAGATTTTATGACATAA 2840
Db 300 CAGCAGAAGGTGCTACTGGTTTGCAGAAAGAGGATTTCAGGAATCGATTTTATGATATAA 359
QY 2841 CTCAATTTCGTAAGSAATCTTGGACTTTCAGCACCTTAATGGATATATTTGAGAGAGAACAGA 2900
Db 360 CTCAGTTTCATAAGAACCTTGGACTTTCAGCACCTTGAATGATATATTTGAGAGAGAACAGA 419
QY 2901 TCACCTTTGGATGTATTTAGTTTGGAGTGGGCGACAGGAGCTGAAGGAGATTGGAATCAATG 2960
Db 420 TTACCTTAGATGTCCTAGTTTGAATGGCCACAAGGAACTGAAAGAGATTGGAATCAATG 479
QY 2961 CTTATGGACATAGGCACAAACTAATTTAAAGGAGTCGAGAGACTTATCTCCGGACAACAG 3020
Db 480 CCTATGGACATCGACACACAGCTGATTAAAGGAGTTTGAAGGCTGATCTCTGGACAACAAG 539
QY 3021 GTCTTAAACCATATTTAACTTTTGAACACCTCTGCTAGTGAACAAATCTTATAGATCTGT 3080
Db 540 GTCTTAACTCCATATTTAACTCTAAACACTCTGCTAGTGAACAAATCTTATAGATCTGT 599
QY 3081 CTCTCTGATGATAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAACTACAGTTCGAGAGC 3140
Db 600 CTCTCTGATGATAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAACTACAGTTCGAGAGC 659
QY 3141 ACAGAGATGGAGTGCATGTCAGGTGGAAATCTTCAACAGATACATAATTTCAAGATTTCAGA 3200
Db 660 ACAGAGATGGTGTGTCACGCGAGCGGCTCTTCAACAGATACACATTTCTCAAGATTTCAGA 719
QY 3201 AGGTTTGTACAGAACTATGGAAGATACACTACCTCCCGG-AGAAAGAGTCTTCTGAA 3259
Db 720 AGGTTTGTACAGAACTATGGAAGATATACACCTCCCGGAGAAAGAGTCTTCTGAA 779
QY 3260 GAAACACCAACCATGC-CAATGAAGAACTGCTATTTTCATGGGCTCT--CCTTTTGTGAA 3315
Db 780 GAAACACCAACCATGCACAAACCAAGGAGTGTATTTTCATGGGCTCTCTTTTGTGAA 839
QY 3316 TGCAATTATCCACAAAGG--CTTTGATGAAGGCATGC 3351
Db 840 TGGGATTATCCATAAAGGGCTTTTGTATGAAGGCATGC 877

RESULT 7
BQ875327
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BQ875327 human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 648)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
```

MEDLINE 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=st2-RC1-BT0623-120  
 200-011-d04&t3=2000-02-12&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 4  
 High quality sequence stop: 648.

# FEATURES

Source  
 1..648  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="BT0623"  
 /dev\_stage="Adult"  
 /note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
 SmaI; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No. 196  
 ,716 - Ludwig Institute for Cancer Research) profiles  
 into the pUC 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."  
 200 a 118 c 157 g 172 t 1 others

Query Match 15.9%; Score 605; DB 12; Length 648;  
 Best Local Similarity 98.7%; Pred. No. 7.1e-137;  
 Matches 620; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1644 TCCTCCAGAGGGTATCTCATTTAGTAAATTCAGAGGCAGACAGACAATTCGCGAAGCTG 1703  
 DB 23 TCCTCAAGAGGGTATCTCATTTAGG--ATTCAGAGGCAGACAGACAATTCGCGAAGCTG 80  
 QY 1704 CAAAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTCAGAGTGCACACTGCA 1763  
 DB 81 CAAAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTCAGAGTGCACACTGCA 140  
 QY 1764 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTCGACGTGGGTATACACAGTGT 1823  
 DB 141 GAGACATTGAAGGGCGTCAGTCTACACCCNTTCATTTTCGAGCTGGGTATACACAGTGT 200  
 QY 1824 CCGTGGTGAATATCTGCTACAGCATGGAGCTGATGCTGATGCTAAAGATAAAGAGGCC 1883  
 DB 201 CCGTGGTGAATATCTGCTACAGCATGGAGCTGATGCTGATGCTAAAGATAAAGAGGCC 260  
 QY 1884 TTGTACCTTTGCACAAATGATGCTTTTATGACATTATGAAGTTGCAGAACTTCCTTTGTTA 1943  
 DB 261 TTGTACCTTTGCACAAATGATGCTTTTATGACATTATGAAGTTGCAGAACTTCCTTTGTTA 320  
 QY 1944 ACATGGAGCAGTAGTTAATGATGCTGATTTATGAAATTTACACTTTACATGAAGCAG 2003  
 DB 321 ACATGGAGCAGTAGTTAATGATGCTGATTTATGAAATTTACACTTTACATGAAGCAG 380  
 QY 2004 CAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGCTGCAGACCCTACCA 2063  
 DB 381 CAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGCTGCAGACCCTACCA 440  
 QY 2064 AAAAAACAGGGATGGAATACTCCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 2123  
 DB 441 AAAAAACAGGGATGGAATACTCCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 500  
 QY 2124 AGATCTGCTTAGGGAGATGACAGCTTTCTAGATGCTGCCAAGAGGGTTGTTAGCCA 2183  
 DB 501 AGATCTGCTTAGGGAGATGACAGCTTTAGCTAGATGCTGCCAAGAGGGTTGTTAGCCA 560  
 QY 2184 GAGTGAAGAAGTGTCTTCTCTCTGATATGTAATTTGCCGCGATACCCCAAGGCAGACATT 2243

DB 561 GAGTGAAGAAGTGTCTTCTCTGATATGTAATTTGCCGCGATACCCCAAGGCAGACATT 620  
 QY 2244 CAACACCTTTACATTTAGCAGCTGGTTA 2271  
 DB 621 CAACACCTTTACATATAGCAGCTGGTTA 648  
 RESULT 8  
 BE220005/c  
 LOCUS BE220005 622 bp mRNA linear EST 03-JUL-2000  
 DEFINITION hv5h03.x1 NCI\_CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3178325 3'  
 similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED  
 ADP-RIBOSE POLYMERASE. ; mRNA sequence.  
 ACCESSION BE220005  
 VERSION BE220005.1 GI:8907323  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 622)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
 info@image.llnl.gov  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 474.  
 Location/Qualifiers  
 1..622  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3178325"  
 /clone\_lib="NCI\_CGAP\_Lu24"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /note="Organ: lung; Vector: pT7T3D-pac (Pharmacia) with a  
 modified polylinker; Plasmid DNA from the normalized  
 library NCI\_CGAP\_Lu5 was prepared, and ss circles were  
 made in vitro. Following HAP purification, this DNA was  
 used as tracer in a subtractive hybridization reaction.  
 The driver was PCR-amplified cDNAs from a pool of 5,000  
 clones made from the same library (cloneIDs  
 1414920-1417991 and 1520904-1522439). Subtraction by Bento  
 Soares and M. Fatima Bonaldo."  
 BASE COUNT 179 a 130 c 128 g 185 t  
 ORIGIN  
 Query Match 15.6%; Score 593.2; DB 10; Length 622;  
 Best Local Similarity 97.1%; Pred. No. 5.4e-134;  
 Matches 604; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
 QY 3164 GGAATCTTCAACAGATACAAATATTCACAGATTCACAAGGTTTGTGAACAGAACTATGG 3223  
 DB 622 GCAATCTTCCACAGTACTATATTCACAGATTCACAGGTTCTGTACACAGTACTATCG 563  
 QY 3224 GAAAGATACACTCACGGGAGAAAAGATTCTCTGAAGAAAACACCAACCATGCAATGAA 3283  
 DB 562 CAAAGATACACTCACGGGAGAAAAGATTCTCTGAAGAAAACCCCAACCATGCAATGAA 503  
 QY 3284 CGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTTGATGAA 3343  
 DB 502 CGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCAAAGGCTTTTGATGAA 443



modified polylinker; plasmid DNA from the normalized library NCI-CGAP\_Lu24 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 179 a 130 c 124 g 183 t  
 ORIGIN

Query Match 15.4%; Score 584; DB 10; Length 616;  
 Best Local Similarity 96.8%; Pred. No. 9.5e-132;  
 Matches 596; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 3344 AGGCATCGTACATAGGTGGTATGTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTTCC 3403  
 DB 442 AGGCATCGTACATAGGTGGTATGTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTTCC 383  
 QY 3404 AAAAGCAATCAATATGATATGTAATGGAATGGAGAGGTACTGGGTCTCCAGTTTCAAAAAGAC 3463  
 DB 382 AAAAGCAATCAATATGATATGTAATGGAATGGAGAGGTACTGGGTCTCCAGTTTCAAAAAGAC 323  
 QY 3464 AGATCTCTTTACATTTGCCACAGCAGCTGCTCTTTTGGCCGGTAACTTTGGGAAAGTCT 3523  
 DB 322 AGATCTCTTTACATTTGCCACAGCAGCTGCTCTTTTGGCCGGTAACTTTGGGAAAGTCT 263  
 QY 3524 TTCTGTCAGTTTCAGTGCATGAAATGCAATCTCTCTCCAGTGCATCACTCAGTCACT 3583  
 DB 262 TTCTGTCAGTTTCAGTGCATGAAATGCAATCTCTCTCCAGTGCATCACTCAGTCACT 203  
 QY 3584 GGTAGGCCAGTGAATAGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643  
 DB 202 GGTAGGCCAGTGAATAGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACAG 143  
 QY 3644 GCTTATCTGAGTATTTAATTACTTACCAGATTATGAGGCTGAGGTATGGTCGATGGA 3703  
 DB 142 GCTTATCTGAGTATTTAATTACTTACCAGATTATGAGGCTGAGGTATGGTCGATGGA 83  
 QY 3704 TAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATATCAAAAGCAGCAGTGGC 3763  
 DB 82 TAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATATCAAAAGCAGCAGTGGC 23  
 QY 3764 CTCTACGTTTACTCTCTTGTCT 3785  
 DB 22 CTCTACGTTTACTCTCTTGTCT 1

RESULT 9  
 BE467629/c  
 LOCUS  
 DEFINITION hz67b11.x1 NCI-CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3213021 3', similar to TR:095271 095271 TRF1-INTERACTING ANKYRIN-RELATED  
 ADP-RIBOSE POLYMERASE. ;, mRNA sequence.  
 ACCESSION BE467629  
 VERSION BE467629.1 GI:9513404  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 616)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 CONTACT Robert Strausberg, Ph.D.  
 EMAIL: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LENL, send email to: [info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 402.  
 Location/Qualifiers  
 1. .616  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3213021"  
 /clone\_lib="NCI-CGAP\_Lu24"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH10B"  
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a

BE910901 601661978F1 NCI-CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3962251 5', mRNA sequence.  
 BE910901  
 BE910901.1 GI:10407885  
 EST.  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 956)

RESULT 10  
 BE910901  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE



please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

#### FEATURES

Location/Qualifiers  
1..702  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="B930018021"  
/clone\_lib="RIKEN full-length enriched, 10 days neonate cerebellum"  
/sex="mixed"  
/tissue\_type="cerebellum"  
/dev\_stage="10 days neonate"  
/lab\_host="DH108"  
/note="Site 1: Sali; Site 2: BamHI: cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'  
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAAATAATGATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

BASE COUNT 202 a 160 c 173 g 167 t

Query Match 15.0%; Score 568; DB 10; Length 702;

Best Local Similarity 89.7%; Pred. No. 8.2e-128;

Matches 621; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

QY 1313 TGTGCTGCTGATCTCCATATCCCAAGAAAGCAATATGTGACCTGCTGAAGAAA 1372

DB 12 TGTGCTGCTGATCT-CATATCCCAAGAAAGCAGATATGTGAACGTGCTCAGAAAG 70

QY 1373 GGAGCAACATCAATGAAAGACTTAAGAAATCTTGACTCTCTGCGCGGCTCTGAG 1432

DB 71 GGAGCAACACCAAGAAAGACTTAAGAAATCTTGACTCTCTGCGCGGCTCTGAG 130

QY 1433 AAGCTCATAATGATGTTGTTGAAGTAGTGTGTAACATGAAGCAAGGTTAATGCTCTG 1492

DB 131 AACGCTCAATGATGTTGTTGAAGTAGTGTGTAACATGAAGCAAGGTTAATGCTCTG 190

QY 1493 GATAATCTTGTCAGACTCTCTACAGAGACTGCATATTTGCTGCTCTACAAACCTGC 1552

DB 191 GATAGCTTGGCAGACGCTCATATACAGAGCTGCACACTGTGCTGACCTGCAGACCTGC 250

QY 1553 CGCCTACTCTGAGCTATGCGGTGATCCTTAACATTTATATCCCTTCAGGGGCTTTACTGCT 1612

DB 251 CGCCTGCTCTGAGCTATGCGGTGATCCTTAACATTTATATCCCTTCAGGGGCTTTACTGCT 310

QY 1613 TTACAGATGGGAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCTCATTTAGGTAAT 1672

DB 311 TTGCAATGGGAAATGAAATGTGAGCAGCTGTCTCAAGAGGGCGGCTCACTAGGTCAAC 370

QY 1673 TCAGAGGCGCAGACAGCAATTCCTGGAAGCTGCAGAGCTGCAGATGTCGAAACTGTAAA 1732

DB 371 TCAGAGGCGGCGCAGACAGCTGCTGGAAGCTGCAGAGCTGCAGATGTCGAAACTGTAAA 430

QY 1733 AAATGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGGTCACTCTACACCA 1792

DB 431 AAACCTGCTGCTTTCAGAGTGTCAACATGCAGAGACATTGAAGGACGCGCAGTCAACCCCA 490

QY 1793 CTTTCATTTTTCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGA 1852

DB 491 CTCACCTTTTCAGCTGGGTATACAGAGTGTCTGTTGTTGAATATCTGCTGCAACATGGA 550

QY 1853 GCTGATGTGCATGCTAAAGATAAAGGAGCGCTTGACCTTTTCACAATGCATGTTCTTAT 1912

DB 551 GCTGATGTTTCATGCTAAAGACAAGAGGAGTCTTGACCTTTTACACAAATGCATGTTCTTAT 610

QY 1913 GGACATTTATGAAGTTGTCAGAACTCTCTTTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1972

DB 611 GGACACTATGAAGTTGTCAGAACTCTCTTTGTCAGAGCATGGAGCAGTAGTCAATGCTGCTGAC 670

QY 1973 TTATGGAATTTTACACCTTTTACATGAAGCAGC 2004

DB 671 TTGTGGAAGTTTACACCTTTTACATGAAGCTGC 702

RESULT 12

BE222311/c

LOCUS

DEFINITION

ADP-RIBOSE POLYMERASE. ;, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 585)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cqabbs-r@mail.nih.gov](mailto:cqabbs-r@mail.nih.gov)

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

[info@image.llnl.gov](mailto:info@image.llnl.gov)

Seq primer: -400P from Gibco

High quality sequence stop: 476.

Location/Qualifiers

1..585

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:316617"

/clone\_lib="NCI-CGAP\_Lu24"

/tissue\_type="carcinoid"

/lab\_host="DH108"

/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; plasmid DNA from the normalized

library NCI-CGAP\_Lu5 was prepared, and ss circles were

made in vitro. Following HAP purification, this DNA was

used as tracer in a subtractive hybridization reaction.

The driver was PCR-amplified cDNAs from a pool of 5,000

clones made from the same library (cloneIDs

1414920-1417991 and 1520904-1522439). Subtraction by Bento

Soares and M. Fatima Bonaldo.

BASE COUNT 166 a 126 c 112 g 176 t

ORIGIN

Query Match 14.8%; Score 562.8; DB 10; Length 585;

Best Local Similarity 98.5%; Pred. No. 1.4e-126;

Matches 576; Conservative 0; Mismatches 7; Indels 2; Gaps 1;



SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 553)
AUTHORS	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,A., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J.G., Blisstein,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE	Endocrine Pancreas Consortium
JOURNAL	Unpublished (2000)
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@iobp.harvard.edu Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu) Seq primer: -40RP from Gbco High quality sequence stop: 429.

```

FEATURES
  source
    1. 553
      Location/Qualifiers
        db_xref="taxon:9606"
        clone_lib="Human insulinoma"
        tissue_type="insulinoma"
        lab_host="DH10B (phage-resistant)"
        note="Organ: pancreas; Vector: pBluescript SK-; Site:1;
        xhoI; Site:2; EcoRI; Constructed with lambda ZAPII system
        (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
        pBluescript SK- by Dr. H. Inoue following the Washington
        University protocol
        (http://genome.wustl.edu/est/lambda_protocol.shtml).
        Please contact Hiroshi Inoue, MD/PhD for further
        information on this library (Metabolism Division, Permutt
        Laboratory, Washington University School of Medicine, Box
        8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
        is a Washington University Pancreas EST project library."
    165 a 107 c 121 q 160 t
BASE COUNT

```

Query Match	14.5%	Score 549;	DB 14;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 3.3e-123;		
Matches 549;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	3241	GAGAAAGAAGTTTCTGAAGAAAACCAACCATGCCAATGAACGAATGCTATTTCATGG	3300	
Db	5	GAGAAAGAAGTTTCTGAAGAAAACCAACCATGCCAATGAACGAATGCTATTTCATGG	64	
QY	3301	GTCTCCTTTTGTAATGCAATTATCCACAAGGCTTTTGATGAAGGCATGCGTACATAGG	3360	
Db	65	GTCTCCTTTTGTAATGCAATTATCCACAAGGCTTTTGATGAAGGCATGCGTACATAGG	124	
QY	3361	TGATATGTTTGAGCTGGCATTTATTTGCTGAAACCTCTCCAAAAGCAATCAATATGT	3420	
Db	125	TGATATGTTTGAGCTGGCATTTATTTGCTGAAACCTCTCCAAAAGCAATCAATATGT	184	
QY	3421	ATATGGAATTGAGGAGGTACTGGGTGCCAGTTTCACAAAGACAGATCTTGTGTACATTGG	3480	
Db	185	ATATGGAATTGAGGAGGTACTGGGTGCCAGTTTCACAAAGACAGATCTTGTGTACATTGG	244	
QY	3481	CCACAGGCAGCTGCTCTTTTGCCGGGTAACTTGGGGAAGCTTTTCCCTGCAGTTTCAGTGC	3540	
Db	245	CCACAGGCAGCTGCTCTTTTGCCGGGTAACTTGGGGAAGCTTTTCCCTGCAGTTTCAGTGC	304	

Qy	3541	AATGAAATGGCACAATTCTCTCTCCAGGTATCACTCACTAGTCACTGGTAGGCCACAGTGAAA	3600
Db	305	AATGAAATGGCACAATTCTCTCTCCAGGTATCACTCACTAGTCACTGGTAGGCCACAGTGAAA	364
Qy	3601	TGGCCTAGCATTACCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT	3660
Db	365	TGGCCTAGCATTACCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT	424
Qy	3661	AATTACTTACCAGATTATGAGGCTTGAAGCTATGCTCGATGGATAAATAGTATTTTAAAG	3720
Db	425	AATTACTTACCAGATTATGAGGCTTGAAGCTATGCTCGATGCATGATAAATAGTATTTTAAAG	484
Qy	3721	AAACTAATTCACGTGAACCTTAAATCATCAAGCAGACAGTGGCCCTCTACGTTTACTCCT	3780
Db	485	AAACTAATTCACGTGAACCTTAAATCATCAAGCAGACAGTGGCCCTCTACGTTTACTCCT	544
Qy	3781	TTGCTGAAA	3789
Db	545	TTGCTGAAA	553
RESULT	15		
LOCUS	BG245969		
DEFINITION	602358786F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4487204 5';	1099 bp	mRNA linear EST 13-FEB-2001
ACCESSION	BG245969		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAMi0331 row: 1 column: 21		
	High quality sequence stop: 698.		

```

high quality sequence stop: 89%.
Location/Qualifiers
1. .1099
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4487204"
/clone_lib="NCI_CGAP_Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      370 a   219 c   289 g   221 t
ORIGIN
Query Match          14.3%; Score 544; DB 12; Length 1099;
Best Local Similarity 88.5%; Pred. No. 7.le-122;
Matches 680; Conservative 0; Mismatches 80; Indels 8; Gaps 8;
Qy 2918 GTTGAGATGGGGCCACAGGAAGCTCAAGGAGATTTGGAATCAATGCTTATGGACATTAGGCAC 2977
1 | TTT ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GTTGAAATGGGCCACAGGAATCTAAGAGAGATTTGGAATCAATGCTTATGGACATTAGGCAC 60

```

Qy	2978	AAACTAAATTAAAGAGCTCAGAGACATTAATCTC-CGGACAAACAAGGTCTTTAACCCATAATTT	30303
Db	61	AAGCTGATTTAAAGGAGTTGAAAGAGTGTAATCTCGTGGACAAACAAGGTCTTTAATCCATAATTT	120
Qy	3037	AACCTTTGAACACCTCTGGTAGTGGAAACAATTTCTTATAGATCTGTCTCCTGATGATAAGA	3096
Db	121	AACCTTAACAACACTCTGGTAGTGGAAACAATTTCTCATAGATCTGTCTCCTGATGATAAGA	180
Qy	3097	GTTTTCAGTCTGTGGAGGAGAGATGCAAAAGTACAGTTTCGAGAGCACAGAGATGGAGGTCA	3156
Db	181	ATTTCAGTCTGTGGAAAGAAGAGATGCAGAGTACTGTGCGGAGACACAGAGATGGTGTCTCA	240
Qy	3157	TGCAGGTGGAAATCTTCAACAGATACAATATTTCTCAAGATTTCAGAAAGTTTGTTAACAGAA	3216
Db	241	CGCAGGCGCGTCTTCAACAGATACACATTTCTCAAGATTTCAGAAAGTTTGTTAACAGAA	300
Qy	3217	ACTATGGGAAAGATACACTCACCGGAGAAAAAGTTTCTTGAAGAAAAACACAAACCATGC	3276
Db	301	ATTGTGGGAAAGATATACACACGGAGAAAAAGTTTCTTGAAGAAAAACACAAACCATGC	360
Qy	3277	CAATGAACGAATGCTATTATTCATGGGCTCCCTTTTGTGAATCCATTTATCCACAAGGCTTT	3336
Db	361	AAACGAAGAAGATTATTTTCATGGGCTCCCTTTTGTGAATCCGATTTATCCATAAGGCGTT	420
Qy	3337	TGATGAAAGGCATCGGTACATAGTGGTATTTTGGAGCTGGCATTTATTTTGTCTGAAAA	3396
Db	421	TGATGAAAGGCATGCATACATAGTGGCATGTATGGAGCTGGAATTTATTTTCTCTCGAA	480
Qy	3397	CTCTTCCAAAAGCAATCATATGTATATGGAAATTTGGAGGAGGACTCGGT-GTCCAGTTTC	3455
Db	481	CTCTTCCAAAAGCAATCAATATGTATGGAAATCGGAGGTGGCACCGGATGGTCCAAATTC	540
Qy	3456	ACAAAGACAGATCTTGTTTACATTTGCCACAGCAGCTGTCTTTTTCGGGGTAACCTTGG	3515
Db	541	ACAAAGACAGATCGTGTACATTTTCTCAGGCAGCTGCT-GTTTGTCTGAGTAACCTTGG	599
Qy	3516	GAAAGTCTTTCTCGAGTTTCAGTGCAGTAATGAAATGGCACATTTCTCTCCAGGTCATCACT	3575
Db	600	GAAAGTCTTTCTTGCAGTTTCAGCGCAATGAAATGGCACATTTCTCTCCCGGGCATCACT	659
Qy	3576	CAGTCACTGGTAGGCCAGTGTAAATGGCTTAGCATTAGCTGGAATATGTTATTATTCACAG	3635
Db	660	CGTCACTGGGCGG-CCAGTGTAAATGG-CTAGCAATTAG-TGAATATG-TATTTACAGAG	715
Qy	3636	GAGACAGCGTATCTCTGAGTATTTAAATTAATCTTACCATGATTTAGAGC	3683
Db	716	GAAACAGG-TTATACTGAAATCTTAATTAATTTTACCGATGGTAGGC	762

Search completed: February 11, 2003, 21:15:20  
Job time : 4968.65 secs







Db 1408 TGCTAAGAAAAGGAGCAAAACATCAATGAAAGAGCTAAAGAAATCTTGACTCCTCTGCAGC 1467  
Qy 1441 TGGCATCTGAGAAAGCTCAATAATCATCTTCTGAAGTAGTGGTGAACATGAAAGCAAGG 1500  
Db 1468 TGGCATCTGAGAAAGCTCAATAATGATGTTGTGAAGTAGTGGTGAACATGAAAGCAAGG 1527  
Qy 1501 TTAATGCTCTGGATAAATCTTGGTTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1560  
Db 1528 TTAATGCTCTGGATAAATCTTGGTTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1587  
Qy 1561 TACAAACCTGCGCGCTACTCTGAGCTATGCGGTGATGCTCAACATTAATCCCTTCAGG 1620  
Db 1588 TACAAACCTGCGCGCTACTCTGAGCTATGCGGTGATGCTCAACATTAATCCCTTCAGG 1647  
Qy 1621 GCTTTACTGCTTTACAGATGGGAATCAAAATGTACAGCAACTCCCTCAAGAGGGTATCT 1680  
Db 1648 GCTTTACTGCTTTACAGATGGGAATGAAAATGTACAGCAACTCCCTCAAGAGGGTATCT 1707  
Qy 1681 CATTAGSTAAATTCAGAGCGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1740  
Db 1708 CATTAGSTAAATTCAGAGCGACAGACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1767  
Qy 1741 AAACGTGTAAGAAAATCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 1800  
Db 1768 AAACGTGTAAGAAAATCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 1827  
Qy 1801 AGTCTACACCACTTCAATTTGACGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGC 1860  
Db 1828 AGTCTACACCACTTCAATTTGACGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGC 1887  
Qy 1861 TACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTAACCTTTGCACAATG 1920  
Db 1888 TACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTAACCTTTGCACAATG 1947  
Qy 1921 CATGCTTTATGGACATTAAGAGTTCGACAATCTTCTTTAAACATGGAGCAGTAGTTA 1980  
Db 1948 CATGCTTTATGGACATTAAGAGTTCGACAATCTTCTTTAAACATGGAGCAGTAGTTA 2007  
Qy 1981 ATGTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGACGAAAAAGAAATATG 2040  
Db 2008 ATGTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGACGAAAAAGAAATATG 2067  
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGACACCTTACCAAAAAAAGCAAGGATGGAA 2100  
Db 2068 AAATTTGCAAACTTCTGCTCCAGCATGGTGACACCTTACCAAAAAAAGCAAGGATGGAA 2127  
Qy 2101 ATACTCCTTTGCGATCTGTTAAAGATGAGATACAGATATTCAGATCTGCTTAGGGGAG 2160  
Db 2128 ATACTCCTTTGCGATCTGTTAAAGATGAGATACAGATATTCAGATCTGCTTAGGGGAG 2187  
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTCTTTAGCCAGAGTGAAGAAGTTGCTTT 2220  
Db 2188 ATGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTCTTTAGCCAGAGTGAAGAAGTTGCTTT 2247  
Qy 2221 CTCCTGATAATGTAATTTGCGCGGATACCCCAAGCGACATTTCAACACCTTTACATTTAG 2280  
Db 2248 CTCCTGATAATGTAATTTGCGCGGATACCCCAAGCGACATTTCAACACCTTTACATTTAG 2307  
Qy 2281 CAGCTGGTTATTAATTAATTTAGAACTTCAGAGTATTTGTTACACACGAGGCTCATCTGA 2340  
Db 2308 CAGCTGGTTATTAATTAATTTAGAACTTCAGAGTATTTGTTACACACGAGGCTCATCTGA 2367  
Qy 2341 ATGCCAAGACAAAGGAGGACTTATCTTTACATAATGAGCATCTTACGGGCGATGTAG 2400  
Db 2368 ATGCCAAGACAAAGGAGGACTTATCTTTACATAATGAGCATCTTACGGGCGATGTAG 2427  
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATATGTCATGTGTCAATGCCACGCAAAATGGGCTT 2460  
Db 2428 ATGTAGCAGCTCTACTAATAAAGTATATGTCATGTGTCAATGCCACGCAAAATGGGCTT 2487  
Qy 2461 TCACACCTTTGCAACCAAGCAGCCCAAGGAGGACGACACACACTTTGTGCTTTGTCAG 2520  
Db 2488 TCACACCTTTGCAACCAAGCAGCCCAAGGAGGACGACACACTTTGTGCTTTGTCAG 2547

Qy 2521 CCCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTAGTTT 2580  
Db 2548 CCCATGGAGCTGACCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTAGTTT 2607  
Qy 2581 CAGCGGATGATGTACGCGCTCTTCTGACAGCAGCATGCCCATCTGCTGTGCCCTCTT 2640  
Db 2608 CAGCGGATGATGTACGCGCTCTTCTGACAGCAGCATGCCCATCTGCTGTGCCCTCTT 2667  
Qy 2641 GTTACAAAGCCTCAAGTGTCTCAATGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCT 2700  
Db 2668 GTTACAAAGCCTCAAGTGTCTCAATGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCT 2727  
Qy 2701 CTTTCAAGTCCATCTAGCCCATCAAGCCTTCTGACAGCAGCAGTCTTGCACAACCTTTATCTG 2760  
Db 2728 CTTTCAAGTCCATCTAGCCCATCAAGCCTTCTGACAGCAGCAGTCTTGCACAACCTTTATCTG 2787  
Qy 2761 GGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTTGAAGTGAACAGAGGCTGCTTCCAGTT 2820  
Db 2788 GGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTTGAAGTGAACAGAGGCTGCTTCCAGTT 2847  
Qy 2821 TGGAGAAAAGAGGCTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880  
Db 2848 TGGAGAAAAGAGGCTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2907  
Qy 2881 GACTTGAGCACCTTAATGGATATATTTTGAGAGAGAAGACAGATCACTTTGGATGTATTAGTTG 2940  
Db 2908 GACTTGAGCACCTTAATGGATATATTTTGAGAGAGAAGACAGATCACTTTGGATGTATTAGTTG 2967  
Qy 2941 AGATGGGACAAAGAGCTGAAGGAGTGAAGTGAATCAATCTTATGGACATATAGGCAACAC 3000  
Db 2968 AGATGGGACAAAGAGCTGAAGGAGTGAAGTGAATCAATCTTATGGACATATAGGCAACAC 3027  
Qy 3001 TAATTAAGSAGTTCGAGAGACTTATCTCCGACAAACAGTCTTAACCCATATTTAACTT 3060  
Db 3028 TAATTAAGSAGTTCGAGAGACTTATCTCCGACAAACAGTCTTAACCCATATTTAACTT 3087  
Qy 3061 TGAACACCTCTGCTAGTGGAAACAATTTCTTATAGATCTGCTCTCTGATGATAAGAGTTTC 3120  
Db 3088 TGAACACCTCTGCTAGTGGAAACAATTTCTTATAGATCTGCTCTCTGATGATAAGAGTTTC 3147  
Qy 3121 AGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCCGAGAGCACAGAGATGAGGTCATGCAAG 3180  
Db 3148 AGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTCCGAGAGCACAGAGATGAGGTCATGCAAG 3207  
Qy 3181 GTGCAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTAT 3240  
Db 3208 GTGCAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTTGTAAACAAGAACTAT 3267  
Qy 3241 GGGAAAGATACACTCACCGGAGAAAAAGAGTTTCTGAAGAAAAACCAACCAATGCCAATG 3300  
Db 3268 GGGAAAGATACACTCACCGGAGAAAAAGAGTTTCTGAAGAAAAACCAACCAATGCCAATG 3327  
Qy 3301 AACGAATGCTATTTTCAATGGGTCTCTCTTTGTGAATGCAATATCCACAAGGCTTTGATG 3360  
Db 3328 AACGAATGCTATTTTCAATGGGTCTCTCTTTGTGAATGCAATATCCACAAGGCTTTGATG 3387  
Qy 3361 AAAGGATGCTGATACATAGGTGATGTTTGGAGCTGGCAATTTATTTTGTGTAAGAACTCTT 3420  
Db 3388 AAAGGATGCTGATACATAGGTGATGTTTGGAGCTGGCAATTTATTTTGTGTAAGAACTCTT 3447  
Qy 3421 CCAAAAGCAATCAATATGTAATGGAATTTGGAGGAGTACTGGGTCTCCAGTTCCAGTTACAAAG 3480  
Db 3448 CCAAAAGCAATCAATATGTAATGGAATTTGGAGGAGTACTGGGTCTCCAGTTCCAGTTACAAAG 3507  
Qy 3481 ACAGATCTTGTACATTTGCCAGGAGCTGCTCTTTTCCGGGGTAACTTTGGGGAAGT 3540  
Db 3508 ACAGATCTTGTACATTTGCCAGGAGCTGCTCTTTTCCGGGGTAACTTTGGGGAAGT 3567  
Qy 3541 CTTTCTCTGCAAGTTCACTGCAATGCAATGCAATCTCTCCAGGTCATCACTCAGTCA 3600  
Db 3568 CTTTCTCTGCAAGTTCACTGCAATGCAATGCAATCTCTCCAGGTCATCACTCAGTCA 3627



Db 1528 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1587  
Qy 1561 TACAACCTCGCGCTACTCTCTGAGCTATGGGTGATGATCTTAAACATATATATCCCTTCAGG 1620  
Db 1588 TACAACCTCGCGCTACTCTCTGAGCTATGGGTGATGATCTTAAACATATATATCCCTTCAGG 1647  
Qy 1621 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCTCCAAAGAGGGTATCT 1680  
Db 1648 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCTCTCCAAAGAGGGTATCT 1707  
Qy 1681 CATTAGGTAAATTCAGAGCGACACAGACAAATTTGCTGGGAAGCTGCAAGGCTGGAGATGTCG 1740  
Db 1708 CATTAGGTAAATTCAGAGCGACACAGACAAATTTGCTGGGAAGCTGCAAGGCTGGAGATGTCG 1767  
Qy 1741 AAACGTGTAATAAACTGTCTACTGTTACAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800  
Db 1768 AAACGTGTAATAAACTGTCTACTGTTACAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1827  
Qy 1801 AGCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGC 1860  
Db 1828 AGCTACACCACTTCATTTTCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGC 1887  
Qy 1861 TACACCATGGAGCTGATGTCATGCTTAAGATAAAGAGGCTTGTACCTTTGCACAAATG 1920  
Db 1888 TACAGCATGGAGCTGATGTCATGCTTAAGATAAAGAGGCTTGTACCTTTGCACAAATG 1947  
Qy 1921 CATGTTCTTATGGACATTAAGCTTGCAGAACTTCTGTTAAACATGAGCAGCTAGCTTA 1980  
Db 1948 CATGTTCTTATGGACATTAAGCTTGCAGAACTTCTGTTAAACATGAGCAGCTAGCTTA 2007  
Qy 1981 ATGTAGCTGATTTATGSAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATG 2040  
Db 2008 ATGTAGCTGATTTATGSAATTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATG 2067  
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTTACCACCAAAACAGGATGGAA 2100  
Db 2068 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTTACCACCAAAACAGGATGGAA 2127  
Qy 2101 ATACTCTTTTGATCTCTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2160  
Db 2128 ATACTCTTTTGATCTCTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2187  
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTGTGCTT 2220  
Db 2188 ATGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTGTGCTT 2247  
Qy 2221 CTCCTGATAATGTAAATTTGCCGCGATACCCAAAGGAGACATTTCAACACCTTTACATTTAG 2280  
Db 2248 CTCCTGATAATGTAAATTTGCCGCGATACCCAAAGGAGACATTTCAACACCTTTACATTTAG 2307  
Qy 2281 CAGCTGGTATAATTAATTTAGAAGTTGCAGAGTATTTCTTACACACGGAGCTGATGTGA 2340  
Db 2308 CAGCTGGTATAATTAATTTAGAAGTTGCAGAGTATTTCTTACACACGGAGCTGATGTGA 2367  
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCGATGTAG 2400  
Db 2368 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCGATGTAG 2427  
Qy 2401 ATGTAGCAGCTCTACTATAAAGTATAATGTGATGTGCAATGCCACGGACAAATGGGCTT 2460  
Db 2428 ATGTAGCAGCTCTACTATAAAGTATAATGTGATGTGCAATGCCACGGACAAATGGGCTT 2487  
Qy 2461 TCACACCTTTTGCACGAAGCAGCCAAAAGGAGGACGAACACGCTTTGTGCTTTGTGCTAG 2520  
Db 2488 TCACACCTTTTGCACGAAGCAGCCAAAAGGAGGACGAACACGCTTTGTGCTTTGTGCTAG 2547  
Qy 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580  
Db 2548 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2607  
Qy 2581 CAGCGGATGATGACGCGCTCTCTGACAGAGCGCATGCCCCCACTGCTGCGCCCTCTT 2640

Db 2608 CAGCGGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTGCGCCTCTT 2667  
Qy 2641 GTTACAAGCCTCAAGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCT 2700  
Db 2668 GTTACAAGCCTCAAGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCT 2727  
Qy 2701 CTTACAGTCCATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGCTTTGACAACTTATCTG 2760  
Db 2728 CTTACAGTCCATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGCTTTGACAACTTATCTG 2787  
Qy 2761 GGAGTCTTTTTCAGAACTGCTCTTCAAGTGTAGTTCAAGTGAACAGAGGGTGTCCAGTT 2820  
Db 2788 GGAGTCTTTTTCAGAACTGCTCTTCAAGTGTAGTTCAAGTGAACAGAGGGTGTCCAGTT 2847  
Qy 2821 TGCAGAAAAAGAGGTTCCAGGAGTAGATTTTACATAAAGTCAATTTGTAAGGAATCTTG 2880  
Db 2848 TGCAGAAAAAGAGGTTCCAGGAGTAGATTTTACATAAAGTCAATTTGTAAGGAATCTTG 2907  
Qy 2881 GACTTACAGCACTTAATGGATATATTTGAGAGAGAACAGATFCACTTTGGATGTATAGTTG 2940  
Db 2908 GACTTACAGCACTTAATGGATATATTTGAGAGAGAACAGATFCACTTTGGATGTATAGTTG 2967  
Qy 2941 AGATGGGCAACAGGAGCTCAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000  
Db 2968 AGATGGGCAACAGGAGCTCAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3027  
Qy 3001 TAATTTAAAGGAGTCGAGAGACTTATCTCGGCAACAAGGCTTTAACCCATATTTAACTT 3060  
Db 3028 TAATTTAAAGGAGTCGAGAGACTTATCTCGGCAACAAGGCTTTAACCCATATTTAACTT 3087  
Qy 3061 TGAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATATAAGAGTTTC 3120  
Db 3088 TGAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCTCTGATGATATAAGAGTTTC 3147  
Qy 3121 AGTCTGTGGAGGAGAGATGCAAGTACAGTTCAGAGCACAGATGAGGTCATGCGAG 3180  
Db 3148 AGTCTGTGGAGGAGAGATGCAAGTACAGTTCAGAGCACAGATGAGGTCATGCGAG 3207  
Qy 3181 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTGAGAGGTTTGTAAACAAGAACTAT 3240  
Db 3208 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTGAGAGGTTTGTAAACAAGAACTAT 3267  
Qy 3241 GGAAAGATACACTCACCGGAGAAAAGTCTTCTGAAGAAAACCAACCATGCCAATG 3300  
Db 3268 GGAAAGATACACTCACCGGAGAAAAGTCTTCTGAAGAAAACCAACCATGCCAATG 3327  
Qy 3301 AACGAATGCTTATTTTCATGGGTCTCCTTTTGTGAATGCAATATCCACAAGAGCTTTGATG 3360  
Db 3328 AACGAATGCTTATTTTCATGGGTCTCCTTTTGTGAATGCAATATCCACAAGAGCTTTGATG 3387  
Qy 3361 AAAGGATGCTTACATAGGTGCTATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 3420  
Db 3388 AAAGGATGCTTACATAGGTGCTATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 3447  
Qy 3421 CCAAAAGCAATCAATATGATATGGAATGGAGAGGACTGGGTGTCCAGTTTCCAAAG 3480  
Db 3448 CCAAAAGCAATCAATATGATATGGAATGGAGAGGACTGGGTGTCCAGTTTCCAAAG 3507  
Qy 3481 ACAGATCTTTGTTACATTTGCCACAGGAGCTGCTCTTTTCCCGGTTAACTTGGGAAGT 3540  
Db 3508 ACAGATCTTTGTTACATTTGCCACAGGAGCTGCTCTTTTCCCGGTTAACTTGGGAAGT 3567  
Qy 3541 CTTTCCCTGCAAGTTCAGTGAATGAAATGCAATTTCTCCAGGTCATCTCACTCAGTCA 3600  
Db 3568 CTTTCCCTGCAAGTTCAGTGAATGAAATGCAATTTCTCCAGGTCATCTCACTCAGTCA 3627  
Qy 3601 CTGCTAGGCCAGTGTAAATGGCTTAGCATTTAGCTGAATATGTTATTTACAGAGGAAAC 3660  
Db 3628 CTGCTAGGCCAGTGTAAATGGCTTAGCATTTAGCTGAATATGTTATTTACAGAGGAAAC 3687  
Qy 3661 AGGCTTATCTGAGTATTTAAATTTACTTTACAGATTTATGAGGCTGAGGATGCTGATG 3720  
Db 3688 AGGCTTATCTGAGTATTTAAATTTACTTTACAGATTTATGAGGCTGAGGATGCTGATG 3747

```

QY 3721 GATAAATAGTTATTTTAAAGAACTAATTCCTACTGAACCTAAATATCAATCAAGCAGCAGTG 3780
|||||
Db 3748 GATAAATAGTTATTTTAAAGAACTAATTCCTACTGAACCTAAATATCAATCAAGCAGCAGTG 3807
|||||

QY 3781 GCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3813
|||||
Db 3808 GCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3840
|||||

RESULT 3
AX062247
LOCUS AX062247 4406 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 106 from Patent WO0100849.
ACCESSION AX062247
VERSION AX062247.1 GI:12540148
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4406)
AUTHORS Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 106 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES
Location/Qualifiers
1..4406
/codon_start=1
/organism="Homo sapiens"
/db_xref="taxon:9606"
<2..3790
/feature_start=1
/feature_end=3790
/feature_note="unnamed protein product"
/feature_id="CAC25062.1"
/feature_db_xref="GI:12540149"
/feature_translation="ELAGALPASAAAFODPDGFGALPPPPRSGRAGSPARGARG
GHGTPDPVTAGSQARALSSPGGLALLAGPGLLRLLALLLAVAAADMSGRRC
AGGAACASAAAEVPEAPARELFECRNGDVERKRLVTPKRVNSRDTRGRKSTPLHF
AAGFRKDVYELQNGANVARDDGLIPLHNACSGFGEVYVNLRLRGADPNARDN
WNYTPLEAAIKGKIDVCTVLQHGAEPTIRNTDGTALDLADPSAKVALTGEYKDE
LLESARNEEMKALLPLPNVCHASDRKSTPLHLAGYNRVKIYVOLLVKGADVH
AKDGDVPLHNACSYGHEVTELLVKGACVANDLHQFTPLHEAASKNRVEVCSLL
LSYGADPILLNHNKSAIDLAFTPOLKERLAYEFKGLSLQAAREADVIRKHLISLE
MYNFHPTHTALHCAASAPYPRKQICELLLRKGANINEKREFLPLHVASKEAH
NDVEVYVYKHEAKNALNDGQTSILHRAAYCGHLQTCRLLSYGDSPNIIISLQFTAL
QMGNEVQQLLEGISLGNSEADRLLEAAKAGDVETVKICTVQSVNCRDIEGRQST
PLHFAAGYNRVSVVYLLQHGADVHAKDGLVPLHNACSYGHYEAELLVKGAVVN
VADLWKFPTLHEAAKGYEICKLLOHGADPTKKNRDGNTPDLVKGDDTDIDLLR
GDAALLDAKGLCLARYKLSPPDNVNRDQGRHSTPLHLAGYNNLEVAEYLLQHG
ADVNDKGLLPLHNAASYGHVQVAAALLIYNACVNATDKWAFPLPHEAAKGRQL
CALLAHAGADPTLKNQEGTPLDLVSADVDVSAALITAMPSPALPSCYQVILVYRSP
GATADALSSPSSSLSAASLNDLSGFSSELSVSSSGTEGASLEKKEVPQVDF
SITQVRNLGLEHMDIFERQIITLDVLVEMGKELKEIGINAYGHRKLLKGVRLI
SGOGLNPLTLNTSGSCTILLDLPDDKRFQSVSEEMQSTVYRHRDGHAGGFNRY
ILLKQKVCNKLWERYTHRRKEVSEENHNHNLPHGSPFVNAILHKGFDERHAY
NIGMGFAGITFYAENSNSQNYVYGGGTGCPVHKDRSCYICHRQLLEFCVTLKGSFL
QFSAMKMAHSPPHSHSVTGRPSVNGLALEAYIYRGEQAYPEYLITYQIMRPEGMVDG
F
BASE COUNT 1268 a 910 c 1054 g 1174 t
ORIGIN
Query Match 99.4%; Score 3791.4; DB 6; Length 4406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1 CGCGCTGCTCCCGCGCGGCGAGCGGCGGCGAGGAGCCAGCGGCGCGCG 60
|||||
Db 67 CGCGCTGCTCCCGCGCGGCGAGCGGCGGCGGAGGAGCCAGCGGCGCGCG 126
|||||
QY 61 TGGGCGCGCCCATGGGACTGCGCGCGGATCCCGGTGCACACAGGAGCCAGCGCGCGG 120
|||||
Db 127 TGGGCGCGG-C-CATGGGACTGCGCGCGGATCCCGGTGCACACAGGAGCCAGCGCGCGG 185
|||||

```

```

QY 121 CCTGAGCGCGCTTCTTCGGGGGGGCGCTCGCCCTCTGCTCGGGGGCGGGGCTCTCTGC 180
|||||
Db 186 CCTGAGCGCGCTTCTTCGGGGGGGCGCTCGCCCTCTGCTCGGGGGCGGGGCTCTCTGC 245
|||||
QY 181 TCGGGTTGCTGGCGCTTGTGCTGGCTGTGGCGGCGGCGGAGGATCATGTTCGGGTTCGGCGCT 240
|||||
Db 246 TCGGGTTGCTGGCGCTTGTGCTGGCTGTGGCGGCGGCGGAGGATCATGTTCGGGTTCGGCGCT 305
|||||
QY 241 GCGCGCGGGGAGCGCGCTCGCGCGAGCGCGCGCGCGGCGGCGGAGCGCGGCGCGCGCC 300
|||||
Db 306 GCGCGCGGGGAGCGCGCTCGCGAGCGCGCGCGCGCGGCGGAGCGCGCGCGCGCGCC 365
|||||
QY 301 GAGAGCTGTTGAGAGCGGTGCGCAACGGGACGCTGGAACAGTCAAGAGGCTGCTGACGC 360
|||||
Db 366 GAGAGCTGTTGAGAGCGGTGCGCAACGGGACGCTGGAACAGTCAAGAGGCTGCTGACGC 425
|||||
QY 361 CTGAGAAGGTGAACAGCGCGGACCGGCGGCGGAGGAAATCCACCCGCTGCACCTTCGCCG 420
|||||
Db 426 CTGAGAAGGTGAACAGCGCGGACCGGCGGCGGAGGAAATCCACCCGCTGCACCTTCGCCG 485
|||||
QY 421 CAGGTTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAGAAATGCTGCAAAATGTCCAAG 480
|||||
Db 486 CAGGTTTTGGCGGAAAGACGCTAGTTGAATATTTGCTTCAGAAATGCTGCAAAATGTCCAAG 545
|||||
QY 481 CACGTGATGATGGGGCGCTTATTCCTCTTCAATATGCATGCTCTTTTGGTTCATGCTGAAG 540
|||||
Db 546 CACGTGATGATGGGGCGCTTATTCCTCTTCAATATGCATGCTCTTTTGGTTCATGCTGAAG 605
|||||
QY 541 TAGTCAATCTCTTTTGGCACATGTTGCGAGACCCCAATGCTCGAGATAATTTGAATATTA 600
|||||
Db 606 TAGTCAATCTCTTTTGGCACATGTTGCGAGACCCCAATGCTCGAGATAATTTGAATATTA 665
|||||
QY 601 CTCCTCTCCATGAAGCTGCAATTAAGAGAAAGATTGATGTTTTCATTTGCTGCTTTACAG 660
|||||
Db 666 CTCCTCTCCATGAAGCTGCAATTAAGAGAAAGATTGATGTTTTCATTTGCTGCTTTACAG 725
|||||
QY 661 ATGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 720
|||||
Db 726 ATGGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 785
|||||
QY 721 CATCTGCCAAACGAGTCTACTGTTGAATATGAAGAAGTGAACCTTTAGAAAGTGCCA 780
|||||
Db 786 CATCTGCCAAACGAGTCTACTGTTGAATATGAAGAAGTGAACCTTTAGAAAGTGCCA 845
|||||
QY 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACATTAAATGTCAACTGCCACG 840
|||||
Db 846 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACATTAAATGTCAACTGCCACG 905
|||||
QY 841 CAAAGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
|||||
Db 906 CAAAGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 965
|||||
QY 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGG 960
|||||
Db 966 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCTCATGCTAAAGATAAAGGTGATCTGG 1025
|||||
QY 961 TACCATTACACAATGCTGTTCTTATGTTGCTTATTAAGTAACTGAACCTTTTGGTCAAGC 1020
|||||
Db 1026 TACCATTACACAATGCTGTTCTTATGTTGCTTATTAAGTAACTGAACCTTTTGGTCAAGC 1085
|||||
QY 1021 ATGGTGCTCTGTAAATGAATGGAATGCTGTCGCAATTTCACTCTTTCATGAGGAGCTT 1080
|||||
Db 1086 ATGGTGCTCTGTAAATGGAATGCTGTCGCAATTTCACTCTTTCATGAGGAGCTT 1145
|||||
QY 1081 CTAAGAACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGCTGAGACCCCAACTGC 1140
|||||
Db 1146 CTAAGAACAGGTTGAAGTATGTTCTCTTCTTAAGTTATGCTGAGACCCCAACTGC 1205
|||||
QY 1141 TCAATTTGCACAAATAAAAGTGTATAGACTTGGCTCCACACACAGTAAAGAAAGAT 1200
|||||
Db 1206 TCAATTTGCACAAATAAAAGTGTATAGACTTGGCTCCACACACAGTAAAGAAAGAT 1265
|||||
QY 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1260
|||||

```







QY 961 TACCATTACACAATGCCCTGTTCTTATGGTCATTATGAAGTAACGTAACCTTTTGGTCAAGC 1020  
DB 1612 TACCATTACACAATGCCCTGTTCTTATGGTCATTATGAAGTAACGTAACCTTTTGGTCAAGC 1671  
QY 1021 ATGGTGCCGTGTGAATGAATGCAATGACCTGTGGCAATTCACCTCTCTTTCATGAGGACGCTT 1080  
DB 1672 ATGGTGCCGTGTGAATGCAATGACCTGTGGCAATTCACCTCTCTTTCATGAGGACGCTT 1731  
QY 1081 CTAGAACAGGGTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACATGTC 1140  
DB 1732 CTAGAACAGGGTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACATGTC 1791  
QY 1141 TCAATTGTCAACAATAAAGTCTATAGACTTGGCTGCCACACACAGCTTAAAGAAAGAT 1200  
DB 1792 TCAATTGTCAACAATAAAGTCTATAGACTTGGCTGCCACACACAGCTTAAAGAAAGAT 1851  
QY 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAAGCTGATGTTACTC 1260  
DB 1852 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGGAAGCTGATGTTACTC 1911  
QY 1261 GAATCAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCATCTCTCAACACATGAAA 1320  
DB 1912 GAATCAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCATCTCTCAACACATGAAA 1971  
QY 1321 CAGCATTTGCTGCTGCTGCTCATCTCCATATCCAAAAGAAAGCAAAATATGTGAAGTGT 1380  
DB 1972 CAGCATTTGCTGCTGCTGCTCATCTCCATATCCAAAAGAAAGCAAAATATGTGAAGTGT 2031  
QY 1381 TGCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAGAAATCTTGACTCCTCTGCAAG 1440  
DB 2032 TGCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAGAAATCTTGACTCCTCTGCAAG 2091  
QY 1441 TGGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACATCAAGCAAGG 1500  
DB 2092 TGGCATCTGAGAAAGCTCATTAATGATGTTGTTGAAGTAGTGGTGAACATCAAGCAAGG 2151  
QY 1501 TTAATGCTCTGATAATCTTGCTGAGACTCTCTACACAGAGCTGCATATTTGTTGCTCATC 1560  
DB 2152 TTAATGCTCTGATAATCTTGCTGAGACTCTCTACACAGAGCTGCATATTTGTTGCTCATC 2211  
QY 1561 TACAACCTGCGCCTACTCTGAGCTATGGGTGTGATCCTTAACATATATCCCTTCAGG 1620  
DB 2212 TACAACCTGCGCCTACTCTGAGCTATGGGTGTGATCCTTAACATATATCCCTTCAGG 2271  
QY 1621 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAGAGGGTATCT 1680  
DB 2272 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAGAGGGTATCT 2331  
QY 1681 CATTAGGTAATTCAGAGGCAGACACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1740  
DB 2332 CATTAGGTAATTCAGAGGCAGACACAAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 2391  
QY 1741 AAACGTAAAAAATCTGTACTGTTAGAGTGTCACTGCAGAGACATTTGAAGGGCGTC 1800  
DB 2392 AAACGTAAAAAATCTGTACTGTTAGAGTGTCACTGCAGAGACATTTGAAGGGCGTC 2451  
QY 1801 AGTCTACACCACTCATTTTGCAGCTGGGTATACAGAGTGTCCGCTGGGAATATCTGC 1860  
DB 2452 AGTCTACACCACTCATTTTGCAGCTGGGTATACAGAGTGTCCGCTGGGAATATCTGC 2511  
QY 1861 TACAGCATGGAGCTGATGTGCATGTTAAAGATAAAGAGGCGCTTGTACCTTTGCAACAATG 1920  
DB 2512 TACAGCATGGAGCTGATGTGCATGTTAAAGATAAAGAGGCGCTTGTACCTTTGCAACAATG 2571  
QY 1921 CATGTTCTTATGGACATATGAAGTTGCAAGACTTCTGTTTAAACATGGAGCAGTAGTTA 1980  
DB 2572 CATGTTCTTATGGACATATGAAGTTGCAAGACTTCTGTTTAAACATGGAGCAGTAGTTA 2631  
QY 1981 ATGTAGCTGATTTATGGAATTTACACCTTTACATGAACAGCAGCAAAAAGAAATATG 2040  
DB 2632 ATGTAGCTGATTTATGGAATTTACACCTTTACATGAACAGCAGCAAAAAGAAATATG 2691

QY 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTCCAGACCTTACCACAAAAAACACAGGATGAA 2100  
DB 2692 AAATTTGCAAACTTCTGCTCCAGCATGGTCCAGACCTTACCACAAAAAACACAGGATGAA 2751  
QY 2101 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2160  
DB 2752 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAG 2811  
QY 2161 ATGCAGCTTTGCTAGATGCTGCCAACAAAGGTTGTTTACGAGAGTGAAGAGTTCTCTT 2220  
DB 2812 ATGCAGCTTTGCTAGATGCTGCCAACAAAGGTTGTTTACGAGAGTGAAGAGTTCTCTT 2871  
QY 2221 CTCTCTGATATGTAATTTGCCGCGATACCCAGGCGAGACATTCAACACCTTTACATTTAG 2280  
DB 2872 CTCTCTGATATGTAATTTGCCGCGATACCCAGGCGAGACATTCAACACCTTTACATTTAG 2931  
QY 2281 CAGCTGGTTATATAATTTAGAAGTTGCAGAGTATTTGTTTACAACACGAGCTGATGTGA 2340  
DB 2932 CAGCTGGTTATATAATTTAGAAGTTGCAGAGTATTTGTTTACAACACGAGCTGATGTGA 2991  
QY 2341 ATGCCAAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400  
DB 2992 ATGCCAAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3051  
QY 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCAGGACAAATGGGCTT 2460  
DB 3052 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCAATGCCAGGACAAATGGGCTT 3111  
QY 2461 TCACACCTTTGCAGGAAGACAGCCCCAAAAGGAGGAGACACAGCTTTGCTTTCTGCTAG 2520  
DB 3112 TCACACCTTTGCAGGAAGACAGCCCCAAAAGGAGGAGACACAGCTTTGCTTTCTGCTAG 3171  
QY 2521 CCATGAGCTGTAGCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580  
DB 3172 CCATGAGCTGTAGCCGACTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3231  
QY 2581 CAGCGGATGATGTAGCGCTCTTCTTGACAGAGCCATGCCCCCATCTGCTCGCCCTCTT 2640  
DB 3232 CAGCAGATGATGTAGCGCTCTTCTTGACAGAGCCATGCCCCCATCTGCTCGCCCTCTT 3291  
QY 2641 GTTACAGCTCAAGTCTCAATGGTGTGAGAGCCAGGAGCCACTGACAGATGCTCTCT 2700  
DB 3292 GTTACAGCTCAAGTCTCAATGGTGTGAGAGCCAGGAGCCACTGACAGATGCTCTCT 3351  
QY 2701 CTTCAGGTCCATCTAGCCCCCATCAAGCCCTTCTGACAGCAGCAGCTTTGACAACCTTATCTG 2760  
DB 3352 CTTCAGGTCCATCTAGCCCCCATCAAGCCCTTCTGACAGCAGCAGCTTTGACAACCTTATCTG 3411  
QY 2761 GGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTGTCTCCAGTT 2820  
DB 3412 GGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGGTGTCTCCAGTT 3471  
QY 2821 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCCTAAGGAATCTTG 2880  
DB 3472 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCCTAAGGAATCTTG 3531  
QY 2881 GACITGAGCAGCTAATGGATATATTTGAGAGAGAACAGATCAGCTTTGGATGTATAGTTG 2940  
DB 3532 GACITGAGCAGCTAATGGATATATTTGAGAGAGAACAGATCAGCTTTGGATGTATAGTTG 3591  
QY 2941 AGATGGGCGACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCGACAAAC 3000  
DB 3592 AGATGGGCGACAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCGACAAAC 3651  
QY 3001 TAATTAAGGAGCTCGAGAGACTTATCTCCGACACAAAGGCTTTAACCCATATTTAACTT 3060  
DB 3652 TAATTAAGGAGCTCGAGAGACTTATCTCCGACACAAAGGCTTTAACCCATATTTAACTT 3711  
QY 3061 TGAACACCTCTGGTAGTGAACAATCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTC 3120  
DB 3712 TGAACACCTCTGGTAGTGAACAATCTTATAGATCTGTCTCTCTGATGATAAAGAGTTTC 3771  
QY 3121 AGTCTGTGGAGAAAGATGCAAAAGTACAGTTTCGAGAGCAGACAGATGGAGGTCATGACG 3180



Db 1262 CTCCTCTCCATGAAGCTCAATTAAAGGARAAGATTGATGTTTGCAATGCTGCTTACAGC 1371  
QY 661 ATGGAGCTGAGCCAAACCACTCCGAATACAGATGGAAGACAGACATGGAATTTAGCAGATC 720  
Db 1322 ATGGAGCTGAGCCAAACCACTCCGAATACAGATGGAAGACAGACATGGAATTTAGCAGATC 1381  
QY 721 CATCTGCCAAAGCAGTGTCTACTGCTGAATATAAGAAAGATGAACCTTTAGAAAGTCCCA 780  
Db 1382 CATCTGCCAAAGCAGTGTCTACTGCTGAATATAAGAAAGATGAACCTTTAGAAAGTCCCA 1441  
QY 781 GGAGTGGCAATGAAGAAATAATGATGCTCTACTACACACCATTAATGCTCAACCTGCCACG 840  
Db 1442 GGAGTGGCAATGAAGAAATAATGATGCTCTACTACACACCATTAATGCTCAACCTGCCACG 1501  
QY 841 CAAGTATGGCAGAAAGTCAACTCCATACATTTGGCAGCAGGATATAACAGAGTAAGA 900  
Db 1502 CAAGTATGGCAGAAAGTCAACTCCATTTACATTTGGCAGCAGGATATAACAGAGTAAGA 1561  
QY 901 TTGTACAGCTGTACTGCAACATGAGCTGATGCTCCATGCTAAAGATAAAGGTGATCTGG 960  
Db 1562 TTGTACAGCTGTACTGCAACATGAGCTGATGCTCCATGCTAAAGATAAAGGTGATCTGG 1621  
QY 961 TACCATTACACAATGCCCTGTTCTTATGGTCATTTAAGTAACCTGAACTTTTGGTCAAGC 1020  
Db 1622 TACCATTACACAATGCCCTGTTCTTATGGTCATTTAAGTAACCTGAACTTTTGGTCAAGC 1681  
QY 1021 ATGGTCCCTGTGTAATGCAATGACATGACATGCGCAATTCACCTCTTTCATGAGCAGCTT 1080  
Db 1682 ATGGTCCCTGTGTAATGCAATGACATGACATGCGCAATTCACCTCTTTCATGAGCAGCTT 1741  
QY 1081 CTAAGAACAGGGTGAAGTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140  
Db 1742 CTAAGAACAGGGTGAAGTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1801  
QY 1141 TCAATTGTGCAATAAAAGTCTATAGACTTTGGCTCCACACCACAGTTTAAAGAAAGAT 1200  
Db 1802 TCAATTGTGCAATAAAAGTCTATAGACTTTGGCTCCACACCACAGTTTAAAGAAAGAT 1861  
QY 1201 TAGCATATGAATTTAAAGGCCACCTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1260  
Db 1862 TAGCATATGAATTTAAAGGCCACCTGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1921  
QY 1261 GAATCAAAAACATCTCTCTCTGGAATGTTGAAATTTTCAAGCATCCCTCAAAACATGAAA 1320  
Db 1922 GAATCAAAAACATCTCTCTCTGGAATGTTGAAATTTTCAAGCATCCCTCAAAACATGAAA 1981  
QY 1321 CAGCATTCGATTTGCTGCTGCATCTCCATATCCCAAAAGAACCAATATGTGAACCTGT 1380  
Db 1982 CAGCATTCGATTTGCTGCTGCATCTCCATATCCCAAAAGAACCAATATGTGAACCTGT 2041  
QY 1381 TGCTAAGAAAGGAGCAACATCAATGAAAGACTAAAGAAATTTCTGACTCCTCTGCACG 1440  
Db 2042 TGCTAAGAAAGGAGCAACATCAATGAAAGACTAAAGAAATTTCTGACTCCTCTGCACG 2101  
QY 1441 TGCCATCTGAGAAAGCTCATATGATGTTGTTGAAAGTGTGAGTGTGAGTGTGAGTGTGAGT 1500  
Db 2102 TGCCATCTGAGAAAGCTCATATGATGTTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2161  
QY 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTTGGTCATC 1560  
Db 2162 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGTTGGTCATC 2221  
QY 1561 TACAACCTGCCGCTACTCTGAGCTATGGTGTGATCTTAAACATATATATATATATATATATAT 1620  
Db 2222 TACAACCTGCCGCTACTCTGAGCTATGGTGTGATCTTAAACATATATATATATATATATATAT 2281  
QY 1621 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCT 1680  
Db 2282 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCAAGAGGGTATCT 2341  
QY 1681 CATTAGGTAATTCAGAGGCACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGTGG 1740

Db 2342 CATTAGGTAATTCAGAGGCACAGACAAATTCCTGGAAGCTGCAAGGCTGGAGATGTGG 2401  
QY 1741 AAAGTGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGCTC 1800  
Db 2402 AAAGTGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGCTC 2461  
QY 1801 AGTCTACACCACCTTCATTTTGGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGC 1860  
Db 2462 AGTCTACACCACCTTCATTTTGGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGC 2521  
QY 1861 TACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTTGACCTTTGCAACAATG 1920  
Db 2522 TACAGCATGGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTTGACCTTTGCAACAATG 2581  
QY 1921 CATGTTCTTATGACACATTAATGAAGTTGCAGAACTTCTTGTAAACATGAGCAGTAGTTA 1980  
Db 2582 CATGTTCTTATGACACATTAATGAAGTTGCAGAACTTCTTGTAAACATGAGCAGTAGTTA 2641  
QY 1981 ATCTAGCTGATTTATGGAATTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAATATG 2040  
Db 2642 ATCTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAATATG 2701  
QY 2041 AAATTTGCAAACTTCCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAAGGATGGAA 2100  
Db 2702 AAATTTGCAAACTTCCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAAGGATGGAA 2761  
QY 2101 ATACTCCTTTGGGATCTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAG 2160  
Db 2762 ATACTCCTTTGGGATCTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTTAGGGGAG 2821  
QY 2161 ATCAGCTTTGCTTAGATGCTGCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAGTGTCTCTT 2220  
Db 2822 ATCAGCTTTGCTTAGATGCTGCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAGTGTCTCTT 2881  
QY 2221 CTCTGTAATGTAATGTCGCGGATACCCAGGACAGATTCACACCTTTACATTTAG 2280  
Db 2882 CTCTGTAATGTAATGTCGCGGATACCCAGGACAGATTCACACCTTTACATTTAG 2941  
QY 2281 CAGCTGGTTAATAATTTAGAAAGTTCAGAGTATTTGTTTACAAACAGGAGCTGATGTA 2340  
Db 2942 CAGCTGGTTAATAATTTAGAAAGTTCAGAGTATTTGTTTACAAACAGGAGCTGATGTA 3001  
QY 2341 ATGCCAAGCAAAAGGAGGACTTATTCCTTTTACATATATGAGCATCTTACGGCATGTAG 2400  
Db 3002 ATGCCAAGCAAAAGGAGGACTTATTCCTTTTACATATGAGCATCTTACGGCATGTAG 3061  
QY 2401 ATGTAGCAGCTCTACTAATTAAGTATTAATGATGTGTAATGCCAGGACAAATGGGCTT 2460  
Db 3062 ATGTAGCAGCTCTACTAATTAAGTATTAATGATGTGTAATGCCAGGACAAATGGGCTT 3121  
QY 2461 TCACACCTTTGCAGAGCAGCCCCAAAAGGAGCAACACAGCTTTGTGCTTTGCTGTAG 2520  
Db 3122 TCACACCTTTGCAGAGCAGCCCCAAAAGGAGCAACACAGCTTTGTGCTTTGCTGTAG 3181  
QY 2521 CCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGACAAACACCTTTAGATTTAGTTT 2580  
Db 3182 CCCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGACAAACACCTTTAGATTTAGTTT 3241  
QY 2581 CAGGGATGATGTACGGCTCTTCTGACACGAGCATGCCCCCATCTGCTCTGCCCTCTT 2640  
Db 3242 CAGCAGATGATGTACGGCTCTTCTGACACGAGCCTGCCCCCATCTGCTCTGCCCTCTT 3301  
QY 2641 GTTACAGCCTCAAGTCTCAATGGTGTGAGAAAGGAGCAACACAGCTGCGAGATGCTCTCT 2700  
Db 3302 GTTACAGCCTCAAGTCTCAATGGTGTGAGAAAGGAGCAACACAGCTGCGAGATGCTCTCT 3361  
QY 2701 CTTCAGCTCCATCTAGCCCATCAAGCTTTCTCAGCCAGCAGTCTTGACAACCTTATCTG 2760  
Db 3362 CTTCAGCTCCATCTAGCCCATCAAGCTTTCTCAGCCAGCAGTCTTGACAACCTTATCTG 3421  
QY 2761 GGAGTTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAAGCAGAGGGTGTCTCCAGTT 2820  
Db 3422 GGAGTTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAAGCAGAGGGTGTCTCCAGTT 3481

```
QY 2821 TGAGAAAAAGGAGTCCAGGAGTAGATTTTACATTAATCGTAAGGAATCTTG 2880
|||||
Db 3482 TGAGAAAAAGGAGTCCAGGAGTAGATTTTACATTAATCGTAAGGAATCTTG 3541

QY 2881 GACTTCAGCACCTAATGGATATATTTTCAGAGAGACAGATCACTTTGGATGTATTAGTTG 2940
|||||
Db 3542 GACTTCAGCACCTAATGGATATATTTTCAGAGAGACAGATCACTTTGGATGTATTAGTTG 3601

QY 2941 AGATGGGCACAAGGAGCTGAAGGAGATGGAATCAATGCTTATGACATAGGCACAAAC 3000
|||||
Db 3602 AGATGGGCACAAGGAGCTGAAGGAGATGGAATCAATGCTTATGACATAGGCACAAAC 3661

QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGACACAAGGCTCTTAACCCATATTTAACTT 3060
|||||
Db 3662 TAATTAAGGAGTCGAGAGACTTATCTCCGACACAAGGCTCTTAACCCATATTTAACTT 3721

QY 3061 TGAACACCTCTGTAGTGGACAAATTTCTATAGATCTGTCTCCTGATGATAAAGATTTTC 3120
|||||
Db 3722 TGAACACCTCTGTAGTGGACAAATTTCTATAGATCTGTCTCCTGATGATAAAGATTTTC 3781

QY 3121 AGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTTCGAGAGCACAGAGATGAGGTCATGCAG 3180
|||||
Db 3782 AGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTTCGAGAGCACAGAGATGAGGTCATGCAG 3841

QY 3181 GTGGAATCTTCAACAGATCAATATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTAT 3240
|||||
Db 3842 GTGGAATCTTCAACAGATCAATATTTCTCAAGATTCAGAAGTTTGTAAACAAGAACTAT 3901

QY 3241 GGAAGATACACTCCCGGAGAAAAGTTTCTGAAGAAAACCAACCACTGCAAAATG 3300
|||||
Db 3902 GGAAGATACACTCCCGGAGAAAAGTTTCTGAAGAAAACCAACCACTGCAAAATG 3961

QY 3301 AACGAATGCTATTTCACTGGTCTCTCTTTGCTGAATGCAATATATCCACAAGGCTTTGATG 3360
|||||
Db 3962 AACGAATGCTATTTCACTGGTCTCTCTTTGCTGAATGCAATATATCCACAAGGCTTTGATG 4021

QY 3361 AAAGCATCGGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAACTCTTT 3420
|||||
Db 4022 AAAGCATCGGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAACTCTTT 4081

QY 3421 CCAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGTCCAGTTCACAAAG 3480
|||||
Db 4082 CCAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGGTCCAGTTCACAAAG 4141

QY 3481 ACAGATCTTTTACATTTCCACAGCAGCTGCTCTTTTCCGGGTAAACCTTTGGGAAAGT 3540
|||||
Db 4142 ACAGATCTTTTACATTTCCACAGCAGCTGCTCTTTTCCGGGTAAACCTTTGGGAAAGT 4201

QY 3541 CTTTCTGCAGTTCAGTGAATGAAATGGACATTTCTCCTCAGGTCACTCAGTCA 3600
|||||
Db 4202 CTTTCTGCAGTTCAGTGAATGAAATGGACATTTCTCCTCAGGTCACTCAGTCA 4261

QY 3601 CTGGTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
|||||
Db 4262 CTGGTAGGCCAGTGTAAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAAC 4321

QY 3661 AGGCTTATCCTGAGTATTTAACTTACCAGATATGAGGCTCAAGGTATGGTCGATG 3720
|||||
Db 4322 AGGCTTATCCTGAGTATTTAACTTACCAGATATGAGGCTCAAGGTATGGTCGATG 4381

QY 3721 GATAAATAGTTTATTAAGAACTAATTCCTGAACTGAACCTTAAATCATCAAGCAGCAGTG 3780
|||||
Db 4382 GATAAATAGTTTATTAAGAACTAATTCCTGAACTGAACCTTAAATCATCAAGCAGCAGTG 4441

QY 3781 GCCTCTACGTTTACTCTCTTTGCTGAAAAA 3813
|||||
Db 4442 GCCTCTACGTTTACTCTCTTTGCTGAAAAA 4474

RESULT 6
AF438201 6189 bp mRNA linear PRI 12-DEC-2001
LOCUS
```

```
DEFINITION Homo sapiens tankyrase II mRNA, complete cds.
ACCESSION AF438201
VERSION AF438201.1 GI:17530294
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 6189)
AUTHORS Yin,Y. and Gelmann,E.P.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-2001) Lombardi Cancer Center, Georgetown
University, 3800 Reservoir Rd, NW, Washington, DC 20007, USA
FEATURES
Location/Qualifiers
1..6189
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q24"
380..3880
/note="telomerase binding protein"
/codon_start=1
/product="tankyrase II"
/protein_id="AAL40795.1"
/db_xref="GI:17530295"
/translation="MSGRRCAGGACAAACAAAEVPAARELFEACRNGDVERVKRLV
TPEKNSRDTAGRKSTPLHFAAGFGKQDVVYLLONGANVOARDGGGLPLHNACSGF
HAEVNNLLRHGADPNARDNNYTPHAAIKCKIDVCIVLLOHGAEPITRTDGRTA
LDLADPSAKAVLTGEYKDELLESARSGNEEKMMALLTPNVNCHASDGRKSTPLHLA
AGNRVKIVQLLTHGADVHAKDGLVPLHNACSYGHVYVELLKHGACVYNAADLW
QPTPLHEAASKNRVEVCSLLSYGADPTLLNCHNSAIDLAPTQPKERLAYEFKHS
LQAAAREADVTRIKKHLSEVMVFKHPOTHTALHCAASAPYKPKOICELLKRGAN
INEKTEFLTPLHVAEKANDVVEVVKHAKVNALDNLGOTSLHRAAYCGHLOTCTV
LLLSYGCDPNIISLOGFTALOMNENVOQLLEGLSLGNSEADRLLEAAKAGDVETV
KKLCTVQSVNCRDIEGRQSTPLHFAAGNRVSVVEYLQHGADVHAKDGLVPLHNA
CSYHVEVALLEGLVHKGAVNVADLWKFPLHFAAAGKYEICKLQLLHGADPTKKNRD
GNTPLDLVKGDDTDIDLLRGAADLDAKKGCLARVKLLSPDNVNCRDTPKRRSTP
LHAAAYNNLEVAEYLQHGADVNAQDKGLIPLHNAASYGHVDAALIKYNACVNA
TDKWAFTPLHEAAQKRTQCALLAHAGADPTLKNOGOTPLDLVSADDSVALLTAAM
PPSAPSCVKPQVLNVRSPGATADALSQSPSSPSLSAASLNLNCSFSLSVVS
SSGTSGASLKPEKVEPVDFSTIQVRNLGLEHLMIDIPEREQITLDLVEMGHKELKE
TGINAYGHRHLKIGRLEISQGLNPLYTLTSGSTLILDLSDPKDFEFOSEEM
GSTVREHRDGGHAGGFNRYNLIKTKVKNCKLWERYTHRRKEYVEEENHNHARMLEF
HGSPPVNAIITHGFDEHAYTGMFGAGIYFAENSQSNQYVIGGGTGCPCVHNRDLS
CYICHRQLLFVRTPGLKFLQFSAMKMAHSPGHHVSTGRPSVNGLALAEAYVIYRGEQ
AYPELITYQIMRPEGMDG"
BASE COUNT 1772 a 1233 c 1386 g 1798 t
ORIGIN
```

```
Query Match 99.4%; Score 3791.4; DB 9; Length 6189;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 CGCGCTGCTCCGCCCGCGGGGCGAGCGGGGCGAGGAGCCCGAGGGCGCGCG 60
|||||
Db 157 CGCGCTGCTCCGCCCGCGGGGCGAGCGGGGCGAGGAGCCCGAGGGCGCGCG 216

QY 61 TGGGCGCGGCCCATGGGACTGCGCGCGGATCCCGGTGACAGCAGGAGCCAGCGCGCGG 120
|||||
Db 217 TGGGCGCGG-CCATGGGACTGCGCGGATCCCGGTGACAGCAGGAGCCAGCGCGCGG 275

QY 121 CCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCTCTGTCGCGGGCGCGGGCTCCTGC 180
|||||
Db 276 CCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCTCTGTCGCGGGCGCGGGCTCCTGC 335

QY 181 TCCGGTGTGCTGGCTGTGCTGGCTGTGCGCGCGCCAGGATCATGTGCGGTCGCGGCT 240
|||||
Db 336 TCCGGTGTGCTGGCTGTGCTGGCTGTGCGCGCGCCAGGATCATGTGCGGTCGCGGCT 395

QY 241 GCGCGCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 300
|||||
Db 396 GCGCGCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 455
```

Qy 301 CAGAGCTGTTTCAGCGCTGCCGAACGGGACGTGGAAACAGTCAAGAGGCTGTTGACGC 360  
Dbb 456 CAGAGCTGTTTCAGCGCTGCCGAACGGGACGTGGAAACAGTCAAGAGGCTGTTGACGC 515  
Qy 361 CTGAGAAAGTGAACAGCGCGGACACGGCGGCGAGGAATCCACCCCGCTGCACCTTGCOCG 420  
Dbb 516 CTGAGAAAGTGAACAGCGCGGACACGGCGGCGAGGAATCCACCCCGCTGCACCTTGCOCG 575  
Qy 421 CAGGTTTTGGCGGGAAGACGCTAGTTGAATATTTGCTTCAGAAATGGTGAATGTTCCAAAG 480  
Dbb 576 CAGGTTTTGGCGGGAAGACGCTAGTTGAATATTTGCTTCAGAAATGGTGAATGTTCCAAAG 635  
Qy 481 CACGTGATGATGGGGGCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540  
Dbb 636 CACGTGATGATGGGGGCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 695  
Qy 541 TAGTCAATCTCCTTTTGGGACATGGTCAGACGCCCAATGCTCGAGATAATTTGGAATTTATA 600  
Dbb 696 TAGTCAATCTCCTTTTGGGACATGGTCAGACGCCCAATGCTCGAGATAATTTGGAATTTATA 755  
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTTGCATTGCTGTTTACAGC 660  
Dbb 756 CTCCTCTCCATGAAGCTGCAATTAAGGAAAGATTGATGTTTGCATTGCTGTTTACAGC 815  
Qy 661 ATGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 720  
Dbb 816 ATGGAGCTGAGCCCAACCATCCGAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 875  
Qy 721 CATCTGCCAAAGCAGTGTCTTACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCCTA 780  
Dbb 876 CATCTGCCAAAGCAGTGTCTTACTGGTGAATATAAGAAAGATGAATCTTTAGAAAGTGCCTA 935  
Qy 781 GGAGTGGCAATCAGAAAAATGATGGCTCTACTCACACCATTAATAATGCTCAACTGCCACG 840  
Dbb 936 GGAGTGGCAATCAGAAAAATGATGGCTCTACTCACACCATTAATAATGCTCAACTGCCACG 995  
Qy 841 CAAATGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1055  
Dbb 996 CAAATGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1055  
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCCATGCTAAAGATTAAGGTGATCTGG 960  
Dbb 1056 TTGTACAGCTGTTACTGCAACATGGAGCTGATGCCATGCTAAAGATTAAGGTGATCTGG 1115  
Qy 961 TACCATTACAAATGCCCTGTTCTTATGCTCATTTAGAGTAAGTGAACCTTTTGGTCAAGC 1020  
Dbb 1116 TACCATTACAAATGCCCTGTTCTTATGCTCATTTAGAGTAAGTGAACCTTTTGGTCAAGC 1175  
Qy 1021 ATGGTGCCTGTGTAATGCAATGGACTTTGGGCAATTCACCTCCTTTCATGAGCAGCTT 1080  
Dbb 1176 ATGGTGCCTGTGTAATGCAATGGACTTTGGGCAATTCACCTCCTTTCATGAGCAGCTT 1235  
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTTAAAGTTATGGTGCAGACCCACACTGC 1140  
Dbb 1236 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTTAAAGTTATGGTGCAGACCCACACTGC 1295  
Qy 1141 TCAATTTGCACAAATAAAGTGTCTAGACTTTGGCTCCACACACACAGTTAAAGAAAGAT 1200  
Dbb 1296 TCAATTTGCACAAATAAAGTGTCTAGACTTTGGCTCCACACACACAGTTAAAGAAAGAT 1355  
Qy 1201 TAGCATATGAATTTAAAGGCCACTGTTTGGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1260  
Dbb 1356 TAGCATATGAATTTAAAGGCCACTGTTTGGCTGCAAGCTGCAGAGAGCTGATGTTACTC 1415  
Qy 1261 GAATCAAAAACATCTCTCTGGAATTTGGTGAATTTCAAGCATCTTCAACACATGAAGA 1320  
Dbb 1416 GAATCAAAAACATCTCTCTGGAATTTGGTGAATTTCAAGCATCTTCAACACATGAAGA 1475  
Qy 1321 CAGCATTTGATTTGCTGCTGATCTCCATATCCCAAAAGAAAGCAAAATATGTTGAACGT 1380  
Dbb 1476 CAGCATTTGATTTGCTGCTGATCTCCATATCCCAAAAGAAAGCAAAATATGTTGAACGT 1535  
Qy 1381 TGCTAAGAAAGGAGCAAAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCTCTGACG 1440

Dbb 1536 TGCTAAGAAAGGAGCAAAACATCAATGAAAGACTTAAACAAATTTCTGACTCTCTGACG 1595  
Qy 1441 TGGCATCTCAGAAAAGCTCATATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAAG 1500  
Dbb 1596 TGGCATCTCAGAAAAGCTCATATATGATGTTTGAAGTAGTGGTGAACATGAAGCAAAG 1655  
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTGCATC 1560  
Dbb 1656 TTAATGCTCTGGATAATCTTGGTCAGACTTCTTACACAGAGCTGCATATTTGGTGCATC 1715  
Qy 1561 TACAAACCTGCCGCCCTACTCTTGCAGCTATGGGTGATCCTTAACATATATATCCCTTCAG 1620  
Dbb 1716 TACAAACCTGCCGCCCTACTCTTGCAGCTATGGGTGATCCTTAACATATATATCCCTTCAG 1775  
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATCTACAGCAACTCTCCCAAGAGGTATCT 1680  
Dbb 1776 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTAAGCAACTCTCCCAAGAGGTATCT 1835  
Qy 1681 CATTAGGTAAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1740  
Dbb 1836 CATTAGGTAAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCG 1895  
Qy 1741 AAACCTGTAAAAAATCTGTACTGTTTTCAGAGTGTCAACTGTCAGAGACATTTGAAGGCGCT 1800  
Dbb 1896 AAACCTGTAAAAAATCTGTACTGTTTTCAGAGTGTCAACTGTCAGAGACATTTGAAGGCGCT 1955  
Qy 1801 AGTCTACACCACCTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGGAATATCTCG 1860  
Dbb 1956 AGTCTACACCACCTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGGAATATCTCG 2015  
Qy 1861 TACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGAGGCGCTTGTACCTTTGCACAATG 1920  
Dbb 2016 TACAGCATGGAGCTGATGTCATGCTTAAAGATAAAGAGGCGCTTGTACCTTTGCACAATG 2075  
Qy 1921 CATGTTCTTATGACATATATGAAGTTGCAGAACTTCTTGTFTAACATGAGCAGTAGTTA 1980  
Dbb 2076 CATGTTCTTATGACATATATGAAGTTGCAGAACTTCTTGTFTAACATGAGCAGTAGTTA 2135  
Qy 1981 ATGCTAGCTGATTTATGGAATTTACACCTTTACATGAACGACGACCAAGAAATATG 2040  
Dbb 2136 ATGCTAGCTGATTTATGGAATTTACACCTTTACATGAACGACGACCAAGAAATATG 2195  
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAACACAGGATGGAA 2100  
Dbb 2196 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAACACAGGATGGAA 2255  
Qy 2101 ATACTCCTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAG 2160  
Dbb 2256 ATACTCCTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAG 2315  
Qy 2161 ATGAGCTTTGCTAGATGCTGCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAACTTGTCTT 2220  
Dbb 2316 ATGAGCTTTGCTAGATGCTGCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAACTTGTCTT 2375  
Qy 2221 CTCTCATATAATGTAATTTGCCGCGATACCAAGCAGACATTCACACCTTTACATTTAG 2280  
Dbb 2376 CTCTCATATAATGTAATTTGCCGCGATACCAAGCAGACATTCACACCTTTACATTTAG 2435  
Qy 2281 CAGCTGTTTATATAATTTAGAAGTTGACAGTATTTGTTTACAAACGAGGAGCTGATGTA 2340  
Dbb 2436 CAGCTGTTTATATAATTTAGAAGTTGACAGTATTTGTTTACAAACGAGGAGCTGATGTA 2495  
Qy 2341 ATGCCCCAAGCAAGGAGGACTTATTCCTTTACATATGACAGATCTTACGGCAGTGTAG 2400  
Dbb 2496 ATGCCCCAAGCAAGGAGGACTTATTCCTTTACATATGACAGATCTTACGGCAGTGTAG 2555  
Qy 2401 ATGTAGCAGCTCTACTATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTT 2460  
Dbb 2556 ATGTAGCAGCTCTACTATAAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTT 2615  
Qy 2461 TCACACCTTTGCGAAGACAGCCCAAGGAGGACACAGCTTTTGTGCTTTGTTGCTAG 2520











/db\_xref="taxon:9606"

/chromosome="10"

/map="10q23.2"

140. .3640

/codon\_start=1

/product="tankyrase 2"

/protein\_id="AAK13463.1"

/db\_xref="GI:13161042"

/translation="MSGRRCAGGAGCAASAAEVEPAARELFEACRNGDVEVRKRLV  
TPKVNSTRDTPKSTPLHFAAGFGKRDVVEYLLQNGANVQARDGGLPLPLHNCSTFG  
HAEVNLLRHGADPNARDNNWYPLHAAIKGKIDVCIVLLOHGAEPTRINTDGRTA  
LDLADPSAKAVLTGEYKDELLESARGNEKMMALLTPLNVNCHASDRKSTPLHLA  
AGNVRYIKVOLLHOGADHAKDGLVPLHNACSYGHVEVTELLVHKGACVNDMLW  
QFTPLHEAASKNVEVCSLLSYSGADPTLLNCHNSAIDLAPTOLKRLAEYEFKGS  
LLQAREADVTRIKHLSLEBNFKPQTHETALHCAASAPYKPKQICEILLRKGAN  
INKEKTEFLPLHASEKANDNDVEVYVHEAKNALDNIGOTSLRHAEVLCGLQTCR  
LLSYGSDPNILISLOGLTALQMNENVOQLLQGTSLGSEADRLLEAAKAGVETV  
KKICTVQSNCRDIEGRQSTPLHFAAGYNRVSYVEYLLOHGDADVIKAKGGVPLHNA  
CSYGHVEAELLVHGAENVADLWKFPLHFAAAGKYEICKLLLOHGDAPTKKRD  
GNTPDLVKGDDTDIDLLRGDAALLDAKKGCLARVKLSSPDNVNCRDGTGRHSTP  
LHLAAGYNLEVAEYLLHOGADYNAQDKGLPLPLHNAASYGHVDVAALLIKYNACVNA  
TDKWAFTPLHEAAKQRTQLCALLAHGADPTLKNQEGOTPLDLVSADDSVALLTAA  
PPSALPSCYKQVNLVRSFGATADALSSGSPSSLSAASLNDLSNLSFSLSSVVS  
SSGTEGASSLEKEVEPVDFSIQFVNLGLEHMDIFEREQITLIDLVLVEMGHKELE  
IGTNAYGHRHLKIGVERLISGOGLNPLYTLNLSGSIILIDLSPDKEFQSVBEREM  
QSTVREHRDGHAGGIFNRYNLIKIOKCNKLLWERYTHRRKEVSENNHINHERMLF  
HGSFPVNAIHKGFEDERHAYIGMFGAGIYFAENSXSKSNQYVYVIGGGTCGPVHKDRS  
CYICHROLLFCFVTLGKRSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVYIRGEQ  
AYPEYLITYQIMPEGMVDG"

BASE COUNT 1120 a 788 c 924 g 983 t

Query Match 97.68; Score 3724.8; DB 9; Length 3815;

Best Local Similarity 99.99; Pred. No. 0;

Matches 3726; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	86	GGATCCGCTGACAGCAGGAGCCAAAGCGCGCGCGCGCTTCGCGCGCGG 145
Db	1	GGATCCGCTGACAGCAGGAGCCAAAGCGCGCGCGCGCTTCGCGCGGGG 60
Qy	146	CTCGCCCTCTGCTCGCGGGCGGGGCTCTGCTCGCGTGTGCGGCTGTGCTGGC 205
Db	61	CTCGCCCTCTGCTCGCGGGCGGGGCTCTGCTCGGTTGTGCGGCTGTGCTGGC 120
Qy	206	TGTGGCGCGCGCAGGATCATGTGGGTGCGGCTGCGCGCGGGGAGCGGCTCGCG 265
Db	121	TGTGGCGCGCGCAGGATCATGTGGGTGCGGCTGCGCGCGGGGAGCGGCTCGCG 180
Qy	266	GAGCGCGCGCGCGAGCGCGTGGAGCGCGCGCGCGCGAGAGCTGTTCGAGGCGTCCGCCAA 325
Db	181	GAGCGCGCGCGCGAGCGCGTGGAGCGCGCGCGCGCGAGAGCTGTTCGAGGCGTCCGCCAA 240
Qy	326	CGGGACGCTGGAACAGCTCAAGAGCTGGTCAAGCGCTGAGAAGTGAACACCGCGGACAC 385
Db	241	CGGGACGCTGGAACAGCTCAAGAGCTGGTCAAGCGCTGAGAAGTGAACACCGCGGACAC 300
Qy	386	GGCGGCGAGGAAATCCACCGCTGCACCTTCGCGCAGGTTTGGCGGGAAGACGCTAGT 445
Db	301	GGCGGCGAGGAAATCCACCGCTGCACCTTCGCGCAGGTTTGGCGGGAAGACGCTAGT 360
Qy	446	TGAATATTTGCTTCAAGATGTTGCAATATGTCACAGCAGCTGATGATGGGGCCCTTATTC 505
Db	361	TGAATATTTGCTTCAAGATGTTGCAATATGTCACAGCAGCTGATGATGGGGCCCTTATTC 420
Qy	506	TCCTTCAATATGATGCTCTTTTGGTCATGCTCAAGTACTCAATCTCTTTTGGCACATGG 565
Db	421	TCCTTCAATATGATGCTCTTTTGGTCATGCTCAAGTACTCAATCTCTTTTGGCACATGG 480
Qy	566	TGCAGACCCCAATGCTCGAGATAATTTGGAATATATCTCTCTCCATGAAGCTGCAATTA 625
Db	481	TGCAGACCCCAATGCTCGAGATAATTTGGAATATATCTCTCTCCATGAAGCTGCAATTA 540
Qy	626	AGGAAAGATTGATTGTGCTGTACAGCATGGAGTGAAGCAACCATCCCGAA 685

Db	541	AGGAAGATTGATGTTGCTGCTGTTACAGCATGGAGTGGAGCAACCATCCGAA 600
Qy	686	TACAGATGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAGCAGTCTTACTGG 745
Db	601	TACAGATGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAGCAGTCTTACTGG 660
Qy	746	TGAATATAAGAAGATGAATCTTAGAAAGTCCAGGAGTGGCAATGAAGAAAAAATGAT 805
Db	661	TGAATATAAGAAGATGAATCTTAGAAAGTCCAGGAGTGGCAATGAAGAAAAAATGAT 720
Qy	806	GGCTCTACTCACACCATTAATATGTCNACTGCCACCGCAAGTGTGGCAGAAAAGTCAACTCC 865
Db	721	GGCTCTACTCACACCATTAATATGTCNACTGCCACCGCAAGTGTGGCAGAAAAGTCAACTCC 780
Qy	866	ATTACATTTGGCAGCAGGATATAACAGAGTAAGATTGTACAGCTGTTACTGCAACATGG 925
Db	781	ATTACATTTGGCAGCAGGATATAACAGAGTAAGATTGTACAGCTGTTACTGCAACATGG 840
Qy	926	AGCTGATGTCATGCTTAAAGATAAAGGTGATCTGTGTACCATTACACAATGCTGTCTTTA 985
Db	841	AGCTGATGTCATGCTTAAAGATAAAGGTGATCTGTGTACCATTACACAATGCTGTCTTTA 900
Qy	986	TGCTCATTTAAGTAACTCAACTTTTGTGAAGCATGTGTGCTGTGTAAATGCAATGA 1045
Db	901	TGCTCATTTAAGTAACTCAACTTTTGTGAAGCATGTGTGCTGTGTAAATGCAATGA 960
Qy	1046	CTTGTGGCAATTCACCTCTCTTCATCAGCAGCTTCTAAGACAGGGTTGAAGTATGTTTC 1105
Db	961	CTTGTGGCAATTCACCTCTCTTCATCAGCAGCTTCTAAGACAGGGTTGAAGTATGTTTC 1020
Qy	1106	TCCTCTCTTAAAGTATGTTGTCAGACCCCAACACTGCTCAATTTGTCAACAATAAAGTCTAT 1165
Db	1021	TCCTCTCTTAAAGTATGTTGTCAGACCCCAACACTGCTCAATTTGTCAACAATAAAGTCTAT 1080
Qy	1166	AGACTTGGCTTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTC 1225
Db	1081	AGACTTGGCTTCCACACACAGTTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTC 1140
Qy	1226	GTTCCTGCAAGCTGCGAGAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGA 1285
Db	1141	GTTCCTGCAAGCTGCGAGAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGA 1200
Qy	1286	AATGGTGAATTTCAAGCATCTCAAAACACATGAAGACAGCATTCGATGCTGCTGCATC 1345
Db	1201	AATGGTGAATTTCAAGCATCTCAAAACACATGAAGACAGCATTCGATGCTGCTGCATC 1260
Qy	1346	TCCATATCCCAAGAAAGCAATATGTGAACCTTGTGCTTGAAGAAAGGAGCAACATCAA 1405
Db	1261	TCCATATCCCAAGAAAGCAATATGTGAACCTTGTGCTTGAAGAAAGGAGCAACATCAA 1320
Qy	1406	TGAAAGACTTAAAGATTTCTTGACTCTCTGCGAGTGGCATCTGAGAAAGCTCATATGA 1465
Db	1321	TGAAAGACTTAAAGATTTCTTGACTCTCTGCGAGTGGCATCTGAGAAAGCTCATATGA 1380
Qy	1466	TGTTGTTGAAGTGTGGTGAACATGAAGCAAAAGGTTAATGCTCTGATATATCTGGTCA 1525
Db	1381	TGTTGTTGAAGTGTGGTGAACATGAAGCAAAAGGTTAATGCTCTGATATATCTGGTCA 1440
Qy	1526	GACTTCTCTACAGAGCTGCATATTTGTGTCATCTACAAACCTCGCGCTACTCTCTGAG 1585
Db	1441	GACTTCTCTACAGAGCTGCATATTTGTGTCATCTACAAACCTCGCGCTACTCTCTGAG 1500
Qy	1586	CTATGGGTGTGCTTAAACATTTATCCCTTCAGGGCTTTTACTGCTTTACAGATGGGAAA 1645
Db	1501	CTATGGGTGTGCTTAAACATTTATCCCTTCAGGGCTTTTACTGCTTTACAGATGGGAAA 1560
Qy	1646	TGAAATGTACAGCAACTCTCCAAAGAGGTATCTCAATAGGTAATTCAGAGGACAGAC 1705
Db	1561	TGAAATGTACAGCAACTCTCCAAAGAGGTATCTCAATAGGTAATTCAGAGGACAGAC 1620
Qy	1706	ACAATTCTGGAAGCTGCAAGGCTGGAGATGTGGAACCTGTAAGAACTGTGTAAGTGT 1765

Db 1621 ACAATTGCTGGAAGCTGCAAAAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTAAGTGT 1680  
QY TCAGAGTGTCAACTGCGACAGACATTTGAAGGCGTCACTGTACACCACTTCAATTTTCAGC 1825  
Db 1681 TCAGAGTGTCAACTGCGACAGACATTTGAAGGCGTCACTGTACACCACTTCAATTTTCAGC 1740  
QY TGGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGC 1885  
Db 1741 TGGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTCATGC 1800  
QY TAAAGATAAAGGAGGCGTGTACCTTTGCACATGCGATGCTTATGAGCATTAAGT 1945  
Db 1801 TAAAGATAAAGGAGGCGTGTACCTTTGCACATGCGATGCTTATGAGCATTAAGT 1860  
QY TGCAGAACTCTTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAATTTAC 2005  
Db 1861 TGCAGAACTCTTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAATTTAC 1920  
QY ACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCA 2065  
Db 1921 ACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCA 1980  
QY TGGTGCAGACCCCTACCAAAAAAACAGGATGGAATACCTCTTGGATCTTTGTTAAAGA 2125  
Db 1981 TGGTGCAGACCCCTACCAAAAAAACAGGATGGAATACCTCTTGGATCTTTGTTAAAGA 2040  
QY TGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCGAGCTTTGCTAGATCTGCCAA 2185  
Db 2041 TGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCGAGCTTTGCTAGATCTGCCAA 2100  
QY GAAGGCTTTGTTAGCCAGAGTGAAGAGTTGCTTCTCTGATTAATGTTAGAGT 2245  
Db 2101 GAAGGCTTTGTTAGCCAGAGTGAAGAGTTGCTTCTCTGATTAATGTTAGAGT 2160  
QY TACCCAAAGGCAGACATTTCAACACCTTTACATTTAGCAGCTGGTTAATAATTTAGAGT 2305  
Db 2161 TACCCAAAGGCAGACATTTCAACACCTTTACATTTAGCAGCTGGTTAATAATTTAGAGT 2220  
QY TGCAGAGTATTGTTTAAACACGAGCTGATGTAATGCCCAGCAAGAGGAGGACTTAT 2365  
Db 2221 TGCAGAGTATTGTTTAAACACGAGCTGATGTAATGCCCAGCAAGAGGAGGACTTAT 2280  
QY TCCCTTTACATAATGCAGCTCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAGTA 2425  
Db 2281 TCCCTTTACATAATGCAGCTCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAGTA 2340  
QY TAATGATGTGTCATGCCACGACAAATGGGCTTTACACCTTTGCACGAAGCAGCCCA 2485  
Db 2341 TAATGATGTGTCATGCCACGACAAATGGGCTTTACACCTTTGCACGAAGCAGCCCA 2400  
QY AAAGGACCAACACAGCTTTGCTTTGTTGCTAGGCCATGGAGCTGACCCGACTTTAA 2545  
Db 2401 AAAGGACCAACACAGCTTTGCTTTGTTGCTAGGCCATGGAGCTGACCCGACTTTAA 2460  
QY AAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTTCT 2605  
Db 2461 AAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGAGATGTCAGCGCTTCT 2520  
QY GACAGAGCCATGCCCCCATCTGCTGCCCCTTCTTACAAAGCCTCAAGTGTCAATGG 2665  
Db 2521 GACAGAGCCATGCCCCCATCTGCTGCCCCTTCTTACAAAGCCTCAAGTGTCAATGG 2580  
QY TGTGAGAACCCAGAGCCACTGCGAGATGCTCTCTTCAGGTCCATTCAGGATCAAG 2725  
Db 2581 TGTGAGAACCCAGAGCCACTGCGAGATGCTCTCTTCAGGTCCATTCAGGATCAAG 2640  
QY CTTTCTGCGAGCAGCAGCTTTGACAACTTATCTGGGAGTTTTCAGAACTGCTCTCAGT 2785  
Db 2641 CTTTCTGCGAGCAGCAGCTTTGACAACTTATCTGGGAGTTTTCAGAACTGCTCTCAGT 2700  
QY AGTTAGTTCAGTGGAAACAGAGGCTTCTTCCAGTTTGGAGAAAGAGGTTCCAGGAGT 2845  
Db 2701 AGTTAGTTCAGTGGAAACAGAGGCTTCTTCCAGTTTGGAGAAAGAGGTTCCAGGAGT 2760

QY 2846 AGATTTTAGCATACTCAATTCGTAAGGAATCTTTGGACTTTCAGCACCTAATGGATATATT 2905  
Db 2761 AGATTTTAGCATACTCAATTCGTAAGGAATCTTTGGACTTTCAGCACCTAATGGATATATT 2820  
QY TGAGAGAAACAGACATCACTTTGGATGTATTAGTTGAGATGGGCACAGGAGCTGAAGGA 2965  
Db 2821 TGAGAGAAACAGACATCACTTTGGATGTATTAGTTGAGATGGGCACAGGAGCTGAAGGA 2880  
QY GATTGGAATCAATGCTTTATGGACATAGGCACAACTAAITTAAGGAGTTCGAGAGCTTAT 3025  
Db 2881 GATTGGAATCAATGCTTTATGGACATAGGCACAACTAAITTAAGGAGTTCGAGAGCTTAT 2940  
QY CTCGGACAAACAGGCTTTAAACCCATATTTAACTTTGAACACCTCTGTTAGTGAACAAT 3085  
Db 2941 CTCGGACAAACAGGCTTTAAACCCATATTTAACTTTGAACACCTCTGTTAGTGAACAAT 3000  
QY TCTTATAGATCTCTCTCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAG 3145  
Db 3001 TCTTATAGATCTCTCTCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAG 3060  
QY TACAGTTTCGAGACACAGATGGAGTCAATGAGGTGGAATCTTCAACAGATACAAATAT 3205  
Db 3061 TACAGTTTCGAGACACAGATGGAGTCAATGAGGTGGAATCTTCAACAGATACAAATAT 3120  
QY TCTCAAGATTCAGAAGTTTGTAAACAAAGAACTATGGAAAGATACACTCACCGGAGAA 3265  
Db 3121 TCTCAAGATTCAGAAGTTTGTAAACAAAGAACTATGGAAAGATACACTCACCGGAGAA 3180  
QY AGAAGTTTCTGAAGAAACCAACCACTGCAATGAACGAATGCTATTTCATGGTCTCC 3325  
Db 3181 AGAAGTTTCTGAAGAAACCAACCACTGCAATGAACGAATGCTATTTCATGGTCTCC 3240  
QY TTTTGTGAATGCAATTTATCCCAAAAGGCTTTGATGAAGGCATGCGTACATAGTGGTAT 3385  
Db 3241 TTTTGTGAATGCAATTTATCCCAAAAGGCTTTGATGAAGGCATGCGTACATAGTGGTAT 3300  
QY GTTTGGAGCTGGCAATTTATTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGTATATGG 3445  
Db 3301 GTTTGGAGCTGGCAATTTATTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGTATATGG 3360  
QY AATTGGAGAGGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTACATTTGCCACAG 3505  
Db 3361 AATTGGAGAGGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTACATTTGCCACAG 3420  
QY GCAGCTGCTTTTGCCTGCTAACTTTGGAAAGTCTTTCCCTGAGTTCAGTGCATGAA 3565  
Db 3421 GCAGCTGCTTTTGCCTGCTAACTTTGGAAAGTCTTTCCCTGAGTTCAGTGCATGAA 3480  
QY AATGGCACATTTCTCTCCAGGTCATCAGTCACTGCTAGGCCAGTGAATGGCCT 3625  
Db 3481 AATGGCACATTTCTCTCCAGGTCATCAGTCACTGCTAGGCCAGTGAATGGCCT 3540  
QY AGCATTTAGCTGAATGTATTATTTACAGAGGAGAACAGGCTTATCTCTGAGTATTAAATAC 3685  
Db 3541 AGCATTTAGCTGAATGTATTATTTACAGAGGAGAACAGGCTTATCTCTGAGTATTAAATAC 3600  
QY TTACCAAGATTATGAGGCTTGAAGGTATGTTTACAGAGGAGAACAGGCTTATCTCTGAGTATTAAATAC 3745  
Db 3601 TTACCAAGATTATGAGGCTTGAAGGTATGTTTACAGAGGAGAACAGGCTTATCTCTGAGTATTAAATAC 3660  
QY ATTCCACTGAACCTAAAATCATCAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3805  
Db 3661 ATTCCACTGAACCTAAAATCATCAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3720  
QY AAAAAAAA 3813  
Db 3721 AAAAAAAA 3728

RESULT 9  
AX062241  
LOCUS

AX062241 4127 bp DNA linear PAT 24-JAN-2001

```
DEFINITION Sequence 100 from Patent WO0100849.
ACCESSION AX062241
VERSION AX062241.1 GI:12540142
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4127)
AUTHORS Christenson, E., Demaggio, A.J., Goldman, P.S. and Mcelligott, D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 100 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES
    source      1..4127
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                <2..3511
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC25060.1"
                /db_xref="GI:12540143"
                /translation="ARMISGRRCAGGAGCAASAAAEVPAARELFEACRNGDVERVK
                RLVTPEKNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANYQARDDGGLPLHNAC
                SPGHAEVNNLLRHGADPNARDNNVTPLEHAAIKGIDVCIVLQHGAEPTIRNDIG
                RTALDLADPSAKAVLTGEYKKDELLESARGNEKEMALLTPLNVNCHASDGRKSTPL
                HLAAGYNRVYKIVOLLQHGADYHAKDKDGLVPLHNACSVGHVEYVELLYKHGCACYNAM
                DLMOFTPLHFAASKNRVESCULLSYGADPTLLNCHNKSALDAPTPQKRELAVEFK
                GHSLQAAAREADVTRIKKLSLEMDVNFHPQTHETALHCAASPYPKRKOICELELLRK
                GANINTEKREFLPLHVASEKHNDFVYVYVHEARVNDLNGOTSLHRAAYCGHLQ
                TCRLLSYGDGPNIIISLOQFTALQMGNEVNVQLOEGISIGNSEADROLLEAAKAGDV
                ETYKLLCTVGSNCRDIEGRSTPLHFAAGYNRVSVVEYLLQHGADYHAKDKGLVPL
                HNACSVGHVEYVELLYKHGCACYNAMDLWKEFTPLHFAAKKVEYICKLLQHGADPTK
                NRDGNTPLDLVKNDGTDIDQLLGRDAALLDAKGCCLARVKKLSSPDNVRCDTRGRH
                STPLHUAAGLVNDEAYEYLLQHGADYHAKDKGLVPLHNACSVGHVEYVELLYKHGCACYNAC
                VNATDWAFPLHFAAKKRTOLCALLLAHAGDPTLNQEGQTPDLVDVADVDVSAALLT
                AAMPSPALSCYKPVQLNGVRSPGATADALSSGSPSSLSAASSLDNISGSELS
                WSSSGTSGASLEKKEVPGVDFSIQFVNLNLEHMDIFEREQITLDVIVEMGHKE
                LKEGINAYGRHUKLKGVERLISGOGLNPLYLTANTSGSTLIDLSPPDKFEFOSVE
                EEMQSTVREHRDGGHAGGFENYINLKIQVCKNKLWERYTHRRKEVSENNHANER
                MLPHGSPVNAIHKGFDERHAYIGMFGAGIYFAENSSKSNQIVYIGGGTGCPVHR
                DRSCYICHRQLFVRVTLGKSLFQFSAMKMAHSPPGHSHSVTGRSVNGLALAEYVYIR
                GEQAYPEYLITYQIMRPEGMVDG"
    CDS         3'UTR      1245 a 813 c 934 g 1135 t
    ORIGIN
Query Match      94.2%; Score 3595.8; DB 6; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 215 GGCCAGGATCATGTGGTCCGCGTGGCGCGGGGGAGCGCGCTGGCGGACGCGCG 274
DB 1 GGCCAGGATCATGTGGTCCGCGTGGCGCGGGGGAGCGCGCTGGCGGACGCGCG 60
QY 275 GGCCAGGCGGTGGAGCGCGCGCGGAGAGCTGTTCGAGCGGTGGCGCAACGGGACGT 334
DB 61 GGCCAGGCGGTGGAGCGCGCGCGGAGAGCTGTTCGAGCGGTGGCGCAACGGGACGT 120
QY 335 GGAACGAGTCAAGAGCGGTGGTGGCGGAGCTGTTCGAGCGGTGGCGCAACGGGACG 394
DB 121 GGAACGAGTCAAGAGCGGTGGTGGCGGAGCTGTTCGAGCGGTGGCGCAACGGGACG 180
QY 395 GAAATCCACCGCGTGCATTCGCGCGAGGTTTGGCGGGAAGACGCTAGTTGAATATT 454
DB 181 GAAATCCACCGCGTGCATTCGCGCGAGGTTTGGCGGGAAGACGCTAGTTGAATATT 240
QY 455 GCTTCAGATGTGCAATGTCCAGACGCTGATGATGGGCGCTTATTCCTCTTCATAA 514
DB 241 GCTTCAGATGTGCAATGTCCAGACGCTGATGATGGGCGCTTATTCCTCTTCATAA 300
QY 515 TGCATGCTCTTTTGGTTCATGCTGAAAGTAGTCAATCTCCTTTTGGGACATGGTGCAGACC 574
|||||
301 TGCATGCTCTTTTGGTTCATGCTGAAAGTAGTCAATCTCCTTTTGGGACATGGTGCAGACC 360
QY 575 CAATGCTCAGATAATTTGAATATATACCTCTCTCCATGAAGCTGCAATTTAAAGGAAAGAT 634
|||||
361 CAATGCTCAGATAATTTGAATATATACCTCTCTCCATGAAGCTGCAATTTAAAGGAAAGAT 420
QY 635 TGATGTTTGCATTTGCTGCTTTACAGCATGGAGCTGAGCCCAACCATCCGAAATACAGATGG 694
|||||
421 TGATGTTTGCATTTGCTGCTTTACAGCATGGAGCTGAGCCCAACCATCCGAAATACAGATGG 480
QY 695 AAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATATAA 754
|||||
481 AAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTCTTACTGGTGAATATAA 540
QY 755 GAAAGATGAACCTTTAGAAAGTCCAGGAGTGGCAATGAAGAAAAAATCATGGCTCTACT 814
|||||
541 GAAAGATGAACCTTTAGAAAGTCCAGGAGTGGCAATGAAGAAAAAATCATGGCTCTACT 600
QY 815 CACACCATTTAAATGTCAACTGCCACGCAAGTGTGTCAGAGTGCAGAAAGTCAACTCCATTACATTT 874
|||||
601 CACACCATTTAAATGTCAACTGCCACGCAAGTGTGTCAGAGTGCAGAAAGTCAACTCCATTACATTT 660
QY 875 GGCGAGGAGTATAACAGAGTAAAGATTGTACAGCTGTTTACTGCAACATGGAGCTGATGT 934
|||||
661 GGCGAGGAGTATAACAGAGTAAAGATTGTACAGCTGTTTACTGCAACATGGAGCTGATGT 720
QY 935 CCATGCTAAGATAAAGGTTTGGTCAAGCATGGTGTCTGTGTAATGCAATGGAGCTTGTGGCA 994
|||||
721 CCATGCTAAGATAAAGGTTTGGTCAAGCATGGTGTCTGTGTAATGCAATGGAGCTTGTGGCA 780
QY 995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGTCTGTGTAATGCAATGGAGCTTGTGGCA 1054
|||||
781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGTCTGTGTAATGCAATGGAGCTTGTGGCA 840
QY 1055 ATTCACTCTCTTCATGAGGAGCTTCTTAAAGACAGGTTTGAAGTATGTCTCTCTCTCT 1114
DB 841 ATTCACTCTCTTCATGAGGAGCTTCTTAAAGACAGGTTTGAAGTATGTCTCTCTCTCT 900
QY 1115 AAGTATGTTGTCAGACCCCAACACTGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCA 1174
DB 901 AAGTATGTTGTCAGACCCCAACACTGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCA 960
QY 1175 TCCACACACACAGTTTAAAGAAAGATTTAGCATATGAATTTAAAGGCGCACTCGTGTGCTGA 1234
DB 961 TCCACACACACAGTTTAAAGAAAGATTTAGCATATGAATTTAAAGGCGCACTCGTGTGCTGA 1020
QY 1235 AGCTGCAGAGAGAGCTGATGTTTACTCGAATCAAAAACATCTCTCTCGAATGCTGAA 1294
DB 1021 AGCTGCAGAGAGAGCTGATGTTTACTCGAATCAAAAACATCTCTCTCGAATGCTGAA 1080
QY 1295 TTTCAAGCATCTCAACACATGAAACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1354
DB 1081 TTTCAAGCATCTCAACACATGAAACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1355 CAAAAGAAAGCAATATGTGAAGTGTGCTTGAAGAAAGGAGCAACATCAATGAAAGAC 1414
DB 1141 CAAAAGAAAGCAATATGTGAAGTGTGCTTGAAGAAAGGAGCAACATCAATGAAAGAC 1200
QY 1415 TAAAGAAATTTCTGACTCTCTGACGCTGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCA 1474
DB 1201 TAAAGAAATTTCTGACTCTCTGACGCTGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCA 1260
QY 1475 AGTAGTGTCAACATGAAGCAAGGTTAATGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1534
DB 1261 AGTAGTGTCAACATGAAGCAAGGTTAATGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1535 ACACAGAGCTGCATATTTGGTGTGCTATCTACAAACCTGCCGCTACTCCTCGAGCTATGGGTG 1594
DB 1321 ACACAGAGCTGCATATTTGGTGTGCTATCTACAAACCTGCCGCTACTCCTCGAGCTATGGGTG 1380
QY 1595 TGATCTTAAACATTTATATCTCTTCCAGGCTTTACTGCTTTTACAGATGGAAATGAAATGT 1654
DB 1381 TGATCTTAAACATTTATATCTCTTCCAGGCTTTACTGCTTTTACAGATGGAAATGAAATGT 1440
```

QY 1655 ACAGCAACTCCTCCAAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTCGT 1714  
DB 1441 ACAGCAACTCCTCCAAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTCGT 1500  
QY 1715 GGAAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAGAACTGTGACTGTTCCAGAGTGT 1774  
DB 1501 GGAAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAGAACTGTGACTGTTCCAGAGTGT 1560  
QY 1775 CAAGTCAGAGACATTGAAGGGCTGAGTCTACACCACTTCATTTGCGAGCTGGGTATAA 1834  
DB 1561 CAAGTCAGAGACATTGAAGGGCTGAGTCTACACCACTTCATTTGCGAGCTGGGTATAA 1620  
QY 1835 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGTGTATGTCATGCTAAAGATAA 1894  
DB 1621 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGTGTATGTCATGCTAAAGATAA 1680  
QY 1895 AGAGGCCCTGTACCTTTGCAACAATGCATGCTTATGACATATTAAGTTCGAGAACT 1954  
DB 1681 AGAGGCCCTGTACCTTTGCAACAATGCATGCTTATGACATATTAAGTTCGAGAACT 1740  
QY 1955 TCTTTGTTAAACATGGAGCAGTACTTAATGTAGCTGATTTATGGAATTTACACCTTTTACA 2014  
DB 1741 TCTTTGTTAAACATGGAGCAGTACTTAATGTAGCTGATTTATGGAATTTACACCTTTTACA 1800  
QY 2015 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074  
DB 1801 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860  
QY 2075 CCTTACCAAAAAACAGGATGGAATACTCCTTTTGGATCTGTTTAAAGATGGAGATAC 2134  
DB 1861 CCTTACCAAAAAACAGGATGGAATACTCCTTTTGGATCTGTTTAAAGATGGAGATAC 1920  
QY 2135 AGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTTG 2194  
DB 1921 AGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGGTTG 1980  
QY 2195 TTTAGCCAGATGAAGAAGTGTCTTCTCCTGATATGTAATTTGCGCGATACCCCAAG 2254  
DB 1981 TTTAGCCAGATGAAGAAGTGTCTTCTCCTGATATGTAATTTGCGCGATACCCCAAG 2040  
QY 2255 CAGACATTCACACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAAAGTTGCCAGAT 2314  
DB 2041 CAGACATTCACACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAAAGTTGCCAGAT 2100  
QY 2315 TTTGTTTACACACGAGCTGATGTGAATGCCCAAGCAAGAGGAGACTTTATTCCTTTACA 2374  
DB 2101 TTTGTTTACACACGAGCTGATGTGAATGCCCAAGCAAGAGGAGACTTTATTCCTTTACA 2160  
QY 2375 TAATGCAGCATCTTAGGGCATGTAGATGTAGCAGCTCTACTAATAAGATATAATGCATG 2434  
DB 2161 TAATGCAGCATCTTAGGGCATGTAGATGTAGCAGCTCTACTAATAAGATATAATGCATG 2220  
QY 2435 TGTCAATGCCAGGACAAATGGGCTTTACACCTTTGCAAGCAAGCAGCCCAAGAGGAGC 2494  
DB 2221 TGTCAATGCCAGGACAAATGGGCTTTACACCTTTGCAAGCAAGCAGCCCAAGAGGAGC 2280  
QY 2495 AACACAGCTTTGTGCTTTGCTAGCCCATGGAGTGCAGCCGACTCTTTAAATATCAGGA 2554  
DB 2281 AACACAGCTTTGTGCTTTGCTAGCCCATGGAGTGCAGCCGACTCTTTAAATATCAGGA 2340  
QY 2555 AGGACAAACACCTTTAGATTTAGTTTACAGCGATGATGTACAGCGCTTCTTCTGACAGCAG 2614  
DB 2341 AGGACAAACACCTTTAGATTTAGTTTACAGCGATGATGTACAGCGCTTCTTCTGACAGCAG 2400  
QY 2615 CATGCCCATCTGCTCTGCCCTCTGTTTACAAAGCCTCAAGTGCATTAAGTGTGAGAG 2674  
DB 2401 CATGCCCATCTGCTCTGCCCTCTGTTTACAAAGCCTCAAGTGCATTAAGTGTGAGAG 2460  
QY 2675 CCCAGAGCCACTGCAGATGCTCTCTTTCAGTCCATCTAGCCCATCAAGCCTTTCTGC 2734  
DB 2461 CCCAGAGCCACTGCAGATGCTCTCTTTCAGTCCATCTAGCCCATCAAGCCTTTCTGC 2520

QY 2735 AGCCAGCAGCTTTCAGCAACTTATCTGGGAGTCTTTCAGAACTGTCTTCACTAGTTAGTTC 2794  
DB 2521 AGCCAGCAGCTTTCAGCAACTTATCTGGGAGTCTTTCAGAACTGTCTTCACTAGTTAGTTC 2580  
QY 2795 AAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAAGGAGTTCCAGAGCTAGATTTAG 2854  
DB 2581 AAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAAGGAGTTCCAGAGCTAGATTTAG 2640  
QY 2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGGACCTTAATGGATATATTTGAGAGAGA 2914  
DB 2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGGACCTTAATGGATATATTTGAGAGAGA 2700  
QY 2915 ACAGATCACTTTTGGATGATTTAGTTGAGATGGGCGCAAGGAGCTGAAGGAGATTTGGAAT 2974  
DB 2701 ACAGATCACTTTTGGATGATTTAGTTGAGATGGGCGCAAGGAGCTGAAGGAGATTTGGAAT 2760  
QY 2975 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGAGCA 3034  
DB 2761 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGAGCA 2820  
QY 3035 ACAAGGCTTTAAACCATATTTAACTTTGAACCTCTGCTAGTGGACAACTTCTTATAGA 3094  
DB 2821 ACAAGGCTTTAAACCATATTTAACTTTGAACCTCTGCTAGTGGACAACTTCTTATAGA 2880  
QY 3095 TCTGCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGC 3154  
DB 2881 TCTGCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGC 2940  
QY 3155 AGAGCACAGATGGAGTTCATCGAGTGGATCTTCAACAGATACAATATCTCAAGAT 3214  
DB 2941 AGAGCACAGATGGAGTTCATCGAGTGGATCTTCAACAGATACAATATCTCAAGAT 3000  
QY 3215 TCAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCCGAGGAAAAAGTTTC 3274  
DB 3001 TCAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCCGAGGAAAAAGTTTC 3060  
QY 3275 TGAAGAAAAACCAACCACTTCCAAAGCAATGCTATTTTCATGGTCTCCTTTTGTGAA 3334  
DB 3061 TGAAGAAAAACCAACCACTTCCAAAGCAATGCTATTTTCATGGTCTCCTTTTGTGAA 3120  
QY 3335 TGCAATATTCACAAAGGCTTTGATCAAGGCAATGCTACATAGTGGTATGTTTGGAGC 3394  
DB 3121 TGCAATATTCACAAAGGCTTTGATCAAGGCAATGCTACATAGTGGTATGTTTGGAGC 3180  
QY 3395 TGGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATGATATGGAATTTGGAG 3454  
DB 3181 TGGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATGATATGGAATTTGGAG 3240  
QY 3455 AGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTACATTTGCCAGGAGCTGCT 3514  
DB 3241 AGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTACATTTGCCAGGAGCTGCT 3300  
QY 3515 CTTTTCGGGGTAACTCTGGGAAAGTCTTCTGCAAGTTCAGTGCATGAAATGGCACA 3574  
DB 3301 CTTTTCGGGGTAACTCTGGGAAAGTCTTCTGCAAGTTCAGTGCATGAAATGGCACA 3360  
QY 3575 TTCTCTCCAGGTCTACCTCAGTCTAGTGGCCAGCTGTAATGGCTTAGCATTTAG 3634  
DB 3361 TTCTCTCCAGGTCTACCTCAGTCTAGTGGCCAGCTGTAATGGCTTAGCATTTAG 3420  
QY 3635 TGAATATGTTATTTACAGAGGAGACAGGCTTATCTCGATATTTAAATTTACATTTACAGAT 3694  
DB 3421 TGAATATGTTATTTACAGAGGAGACAGGCTTATCTCGATATTTAAATTTACATTTACAGAT 3480  
QY 3695 TATGAGGCTGAAGGTATGTCGATGATGAAATAGTATTTTAAAGAACTAATTTCCACATG 3754  
DB 3481 TATGAGGCTGAAGGTATGTCGATGATGAAATAGTATTTTAAAGAACTAATTTCCACATG 3540  
QY 3755 AACCTAAAAATCATCAAGCAGCAGTGGCTCTACGTTTTTACTCCTTTGCTGAAAAAAA 3813  
DB 3541 AACCTAAAAATCATCAAGCAGCAGTGGCTCTACGTTTTTACTCCTTTGCTGAAAAAAA 3599



Db	1688	GATAAAGRRGSCCTTGACCTTTGGACAATGCAATGATCTTATGGACATATATCAAGTTGCA	174
Qy	1950	GAACTCTCTGTTAAACATGAGCAGTAGTTAATGTAGCTGATGATTTATGGAATTTACACCT	2009
Db	1748	GAACTCTCTGTTAAACATGAGCAGTAGTTAATGTAGCTGATGATTTATGGAATTTACACCT	1807
Qy	2010	TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCCTCTCCAGCATGGT	2069
Db	1808	TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCCTCTCCAGCATGGT	1867
Qy	2070	GCAGACCCTACCAAAAAACACGGGATGAAATACTCCTTTGGATCTTGTAAAGATGGA	2129
Db	1868	GCAGACCCTACCAAAAAACACGGGATGAAATACTCCTTTGGATCTTGTAAAGATGGA	1927
Qy	2130	GATACAGATATTCAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGAGCTGCCAAGAAG	2189
Db	1928	GATACAGATATTCAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGAGCTGCCAAGAAG	1987
Qy	2190	GGTTGTTTACCGAGAGTGAAGAGTTGCTCTCTGATATGATGAAATGCGCGGATACC	2249
Db	1988	AGTTGTTTACCGAGAGTGAAGAGTTGCTCTCTGATATGATGAAATGCGCGGATACC	2047
Qy	2250	CAAGGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA	2309
Db	2048	CAAGGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA	2107
Qy	2310	GAGTATTTGTTACAACACGGAGCTGATGAAATGCCAAGACAAGGAGGACTTATTCCT	2369
Db	2108	GAGTATTTGTTACAACACGGAGCTGATGAAATGCCAAGACAAGGAGGACTTATTCCT	2167
Qy	2370	TTACATAATGCAGCATCTTACGGCATGTAGATGTAGCAGCTCTACTAATAAGATATAAT	2429
Db	2168	TTACATAATGCAGCATCTTACGGCATGTAGATGTAGCAGCTCTACTAATAAGATATAAT	2227
Qy	2430	GCATGTCTAAATGCCAGGACAAATGGGCTTTCACACCTTTGCACGAGCAGCCCAAAG	2489
Db	2228	GCATGTCTAAATGCCAGGACAAATGGGCTTTCACACCTTTGCACGAGCAGCCCAAAG	2287
Qy	2490	GGACGAACACAGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCGGACTCTTAAAAAT	2549
Db	2288	GGACGAACACAGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCGGACTCTTAAAAAT	2347
Qy	2550	CAGGAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGGCTTCTCTGACA	2609
Db	2348	CAGGAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGGCTTCTCTGACA	2407
Qy	2610	GCAGCCATGCCCATCTGCTCTGCCCTTGTTCACAGCCTCAAGTCTCAATGGTGTG	2669
Db	2408	GCAGCCATGCCCATCTGCTCTGCCCTTGTTCACAGCCTCAAGTCTCAATGGTGTG	2467
Qy	2670	AGAAGCCAGGAGCCACTGCAGATGCTCTCTTCAGTTCATCTAGCCCATCAAGCCTT	2729
Db	2468	AGAAGCCAGGAGCCACTGCAGATGCTCTCTTCAGTTCATCTAGCCCATCAAGCCTT	2527
Qy	2730	TCCTGCAGCCAGCAGTCTTCACAACTTATCTGGAGCTTTTTCAGAACCTCTCTTCAGTGT	2789
Db	2528	TCCTGCAGCCAGCAGTCTTCAGAACTTATCTGGAGTTTTCAGAACCTCTCTTCAGTGT	2587
Qy	2790	AGTTCAAGTGGAAACAGAGGCTCTCCAGTTTGGAGAAAAAGAGGTTCCAGGAGTAGAT	2849
Db	2588	AGTTCAAGTGGAAACAGAGGCTCTCCAGTTTGGAGAAAAAGAGGTTCCAGGAGTAGAT	2647
Qy	2850	TTTATGATTAACCTAATTCCTAAGGAATCTTGGACTTGAGCACCCTAATGGATATTTTGAG	2909
Db	2648	TTTATGATTAACCTAATTCCTAAGGAATCTTGGACTTGAGCACCCTAATGGATATTTTGAG	2707
Qy	2910	AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAACAAGGAGCTGAAGAGATT	2969
Db	2708	AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAACAAGGAGCTGAAGAGATT	2767
Qy	2970	GGAAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATCTCC	3029
Db	2768	GGAAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATCTCC	2827

Qy	3030	GGACAACAAGGCTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAAACAATTTCTT	3089
Db	2828	GGACAACAAGGCTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAAACAATTTCTT	2887
Qy	3090	ATAGATCTGTCTCCTGATGATGAAGAAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACA	3149
Db	2888	ATAGATCTGTCTCCTGATGATGAAGAAGTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACA	2947
Qy	3150	GTTTCGAGAGCAGAGATGGAGGTCAATGAGGTGCAAACTTTCAACAGATACAATATTTCTC	3209
Db	2948	GTTTCGAGAGCAGAGATGGAGGTCAATGAGGTGCAAACTTTCAACAGATACAATATTTCTC	3007
Qy	3210	AAGATTCAGAAGGTTTGTTAACAGAAGCAATATGCGGAAAAGATACACTACCCGGAGAAAAGAA	3269
Db	3008	AAGATTCAGAAGGTTTGTTAACAGAAGCAATATGCGGAAAAGATACACTACCCGGAGAAAAGAA	3067
Qy	3270	GTTTTCGAAGAAAACACACCAACCATGCCAATGAACCAATGCTATTTTCATGGGTCTCCCTTTT	3329
Db	3068	GTTTTCGAAGAAAACCAACCAACCATGCCAATGAACCAATGCTATTTTCATGGGTCTCCCTTTT	3127
Qy	3330	GTGAATGCAATTTATCCACAAGGCTTTGATGAAGGCGATGCGTACATAGTGGTATGTTT	3389
Db	3128	GTGAATGCAATTTATCCACAAGGCTTTGATGAAGGCGATGCGTACATAGTGGTATGTTT	3187
Qy	3390	GGAGCTGGCAATTTATTTTCTGCTGAACACTCTTCCAAAAGCAATCAATATGTATGGAATT	3449
Db	3188	GGAGCTGGCAATTTATTTTCTGCTGAACACTCTTCCAAAAGCAATCAATATGTATGGAATT	3247
Qy	3450	GGAGGAGGTACTGGGTGTCACAGTTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG	3509
Db	3248	GGAGGAGGTACTGGGTGTCACAGTTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG	3307
Qy	3510	CTGCTCTTTTCCGGGTAACTTTGGAAAGCTTTTCCTGCAAGTTCAGTGCATGAAGAAATG	3569
Db	3308	CTGCTCTTTTCCGGGTAACTTTGGAAAGCTTTTCCTGCAAGTTCAGTGCATGAAGAAATG	3367
Qy	3570	GCACATTCCTCCAGGTCAATCAGTCAGTCAGTGTAGGCCAGTGTAAATGGCCTAGCA	3629
Db	3368	GCACATTCCTCCAGGTCAATCAGTCAGTCAGTGTAGGCCAGTGTAAATGGCCTAGCA	3427
Qy	3630	TTAGCTGAATATGTTATTTTACAGAGGAGACAGGCTTATCCTGAGTATTTAAATTAATCTTAC	3689
Db	3428	TTAGCTGAATATGTTATTTTACAGAGGAGACAGGCTTATCCTGAGTATTTAAATTAATCTTAC	3487
Qy	3690	CAGATTATGAGGCCTCAAGGTATGTCGATGGATAAATAGTTATTTTAAAGAACTAATTC	3749
Db	3488	CAGATTATGAGGCCTCAAGGTATGTCGATGGATAAATAGTTATTTTAAAGAACTAATTC	3547
Qy	3750	CAGTGAACCTTAAATCATCAAGCAGCAGTGGCCTCTACGTTTTTACTCTCTTGTCTGAAAA	3809
Db	3548	CAGTGAACCTTAAATCATCAAGCAGCAGTGGCCTCTACGTTTTTACTCTCTTGTCTGAAAA	3607
Qy	3810	AAAA 3813	
Db	3608	AAAA 3611	
RESULT	11		
LOCUS	AF264912		
DEFINITION	Homo sapiens tankyrase-like protein (TNKL) mRNA, complete cds.	5810 bp	linear
ACCESSION	AF264912		
VERSION	AF264912.1	GI:12005975	
KEYWORDS	.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
	1 (bases 1 to 3810)		
	Kulmov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K.,		
	Scanlan, M.J., Jongeneel, C.V., Lagarkova, M.A. and Nedospasov, S.A.		
	Cloning and characterization of TNKL, a member of tankyrase gene		
TITLE			

family  
 Genes Immun. 2 (1), 52-55 (2001)  
 MEDLINE 21190090  
 PUBMED 11294570  
 REFERENCE 2 (bases 1 to 5810)  
 AUTHORS Kulimov, A.N., Kuprash, D.V., Petrov, V.N., Vdovichenko, K.K., Kashkarova, O.A., Boichenko, V.E., Nedospasov, S.A. and Lagarkova, M.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-MAY-2000) Laboratory of Molecular Immunology, A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia  
 FEATURES  
 source Location/Qualifiers  
 1. 5810  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10q22-10q24"  
 /tissue\_type="breast carcinoma"  
 /note="derived from clones MO-BC-203 and I203"  
 <1. 5810  
 /gene="TNK1"  
 /note="a new member of tankyrase gene family; current UniGene cluster Hs.168491"  
 1. 3501  
 /gene="TNK1"  
 /note="a new member of tankyrase family"  
 /codon\_start=1  
 /product="tankyrase-like protein"  
 /protein\_id="AAG44694.1"  
 /db\_xref="GI:12005976"  
 /translation="MSGRCAGGAGCAACAAAEVPAARELFEACRNGDVERKRLV  
 TPEVNSRDYAGKSTPLHFAAGFGRDKVDVYLLQNGANVOARDGGLIPLHNACSF  
 HAEVNLRLRHGADPNARDNNWTPLHEAAIKGKIDVCIVLLOHGAEPTRINDGRTA  
 LDLDAPSAKAVLGEYKKDELSEARSNEEKMAILLPLNVNCHASDRKSTPLHLA  
 AGYNRVIAQLLOHGADVHAKDGLVPLHNACSYGHVEYTELLVHKGACVNMDLW  
 QFTPLHFAASLNQVGVCSLLSYGADPTLLNCHNSAIDIAPTPOLKRLAYEFKHS  
 LLOAREADVTRIKHLSLEWVNFKHPOTHEALHCAASPYPKRKQICELLRKGAN  
 INEXTKEFLPLHVAESHKANDVVEVVVHKAVALNDNLGOTSLEHRAAYCGHLQCR  
 LLSYSGDPNLSIQGFTALOMGNENYQQLQEGISLGNSEADROLLEAAKAGDVETV  
 KKLCTGVNSCRDIEGRQSTPLHFAAGFGRDKVDVYLLQNGADVHAKDGLVPLHNA  
 CSYGHVEAEVLLKHGAVNVADLWKFPTLHEAAKGYEICKLLLOHGAEPTRKNRD  
 GNTPLDVLQGDYDIQDLRLGDAALDAKKGCLARVKLSLSPDNVNCRDTOGRHSTP  
 LHLAAGYNVLEAYLLQHGADVNAQDGGILPLHNAASYGHVDVAALLIKYNACVNA  
 TDKWAFTPLHEAAKQGTOLCALILAHGADPTLNKNOGQPLDLVSADVYSALLTAAM  
 PPSALPCYKPOVLNVRSPCATADALLSSGPSPLSLAASLNDLNGSPSELSSVVS  
 SSGTEGASSLEKEVPQVDPSITQFVNRNLGLEHLMDFEREQITLDVLVEMGHKELKE  
 IGINAYGHRHKLINGVBRLLISGQGLNPYLTLNTSGSGLILIDLSPQDFQSVVEEM  
 QSTVREHRRDGHAGGIFNRYNLIKIQKCNKWLERYTHRRKEVSEENHNHANEMLF  
 HGSFPVNAIIHKGFDERHAYIGGMFGAGIYFAENSNSKNOYVYIGGGTGPCVHKDRS  
 CYICHRQLLCRVTLGRKSLQFSAMKMAHSPPGHHSVTGRPSVNLALAEYVIYRGEQ  
 AYPEYLITYQIMRPEGMVDG"  
 complement(1132..1277)  
 /note="SHGC-53837"  
 complement(3665..3975)  
 /note="WI-16054"  
 complement(5544..5777)  
 /note="WI-8339"  
 BASE COUNT 1736 a 1094 c 1231 g 1749 t  
 ORIGIN

Query Match 94.0%; Score 3585.8; DB 9; Length 5810;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 225 ATGTGCGGTCCGCGTGGCGGGGGAGGCGGCTCGCGGAGCGCGCGCGCGAGGCC 284  
 Db 1 ATGTGCGGTCCGCGTGGCGGGGGAGGCGGCTCGCGGAGCGCGCGCGCGAGGCC 60  
 Qy 285 GTGGAGCGCGCGCGAGAGCTGTTTCGAGGCGTGCACCAACGGGAGCTGGAACGAGTC 344  
 Db 61 GTGGAGCGCGCGCGAGAGCTGTTTCGAGGCGTGCACCAACGGGAGCTGGAACGAGTC 120

Qy 345 AAGAGGCTGGTGACGCTGAGAAGGTGAACAGCCGGGACACGCGGCGAGGAATCCACC 404  
 Db 121 AAGAGGCTGGTGACGCTGAGAAGGTGAACAGCCGGGACACGCGGCGAGGAATCCACC 180  
 Qy 405 CCCTGCACTTCGCGCAGGTTTGGCGGAAAGAGAGTGTGAATATTTGCTTCAGAAAT 464  
 Db 181 CCCTGCACTTCGCGCAGGTTTGGCGGAAAGAGAGTGTGAATATTTGCTTCAGAAAT 240  
 Qy 465 GGTGCAATGTCAAGCAGCTGATGATGGGGCTTATTCTCTCTTATCATATGATGCTCT 524  
 Db 241 GGTGCAATGTCAAGCAGCTGATGATGGGGCTTATTCTCTCTTATCATATGATGCTCT 300  
 Qy 525 TTTGGTCATGCTCAAGTAGTCAATCTCTTTTGGCAGATGCTGACAGACCCCAATGCTGA 584  
 Db 301 TTTGGTCATGCTCAAGTAGTCAATCTCTTTTGGCAGATGCTGACAGACCCCAATGCTGA 360  
 Qy 585 GATAATTTGGAATTTATCTCTCTCCATGAAGCTGCAATTTAAAGGAAGATTTGATTTGC 644  
 Db 361 GATAATTTGGAATTTATCTCTCTCCATGAAGCTGCAATTTAAAGGAAGATTTGATTTGC 420  
 Qy 645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAAACCATCCGAATACAGATGGAAGCAGCA 704  
 Db 421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAAACCATCCGAATACAGATGGAAGCAGCA 480  
 Qy 705 TTGGATTTAGCAGATCCATCTGCCAAGCAGTCTTACTTGGTGAATATTAAGAAAGATGA 764  
 Db 481 TTGGATTTAGCAGATCCATCTGCCAAGCAGTCTTACTTGGTGAATATTAAGAAAGATGA 540  
 Qy 765 CTCTTAGAAGTGGCAGGAGTGGCAATGAAGAAATAATGATGGCTCTACTCACACCATTA 824  
 Db 541 CTCTTAGAAGTGGCAGGAGTGGCAATGAAGAAATAATGATGGCTCTACTCACACCATTA 600  
 Qy 825 AATGTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884  
 Db 601 AATGTCAACTGCCACGCAAGTGTGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660  
 Qy 885 TATAACAGAGTAAAGATTGTACAGCTGTTTACTTCAACATGGAGCTGATGTCCATGCTAA 944  
 Db 661 TATAACAGAGTAAAGATTGTACAGCTGTTTACTTCAACATGGAGCTGATGTCCATGCTAA 720  
 Qy 945 GATAAAGGTGATGTGTACCATTTACACAATGCCTGTTCTTTATGGTCAATTTGAAGTAACT 1004  
 Db 721 GATAAAGGTGATGTGTACCATTTACACAATGCCTGTTCTTTATGGTCAATTTGAAGTAACT 780  
 Qy 1005 GAACCTTTTGTCAAGCATGGTGTGTGTAATGCAATGACCTTTGGGCAATTTCACTCCT 1064  
 Db 781 GAACCTTTTGTCAAGCATGGTGTGTGTAATGCAATGACCTTTGGGCAATTTCACTCCT 840  
 Qy 1065 CTTTCATGAGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTTAAAGTTATGTT 1124  
 Db 841 CTTTCATGAGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTCTTAAAGTTATGTT 900  
 Qy 1125 GCAGACCCACACTGCTCAATTTGTACCAATTTAAAGTGTCTATAGCTTGGCTCCACACCA 1184  
 Db 901 GCAGACCCACACTGCTCAATTTGTACCAATTTAAAGTGTCTATAGCTTGGCTCCACACCA 960  
 Qy 1185 CAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCGACCTGTTGCTGCAAGCTGCACGA 1244  
 Db 961 CAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCGACCTGTTGCTGCAAGCTGCACGA 1020  
 Qy 1245 GAAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTCTGGAATGTTGAATTTCAAGCAT 1304  
 Db 1021 GAAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTCTGGAATGTTGAATTTCAAGCAT 1080  
 Qy 1305 CCTCAACACATGAACAGCATTTGCATTTGCTGCTGCTCATCTCCATATCCCAAAAGAAAG 1364  
 Db 1081 CCTCAACACATGAACAGCATTTGCATTTGCTGCTGCTCATCTCCATATCCCAAAAGAAAG 1140  
 Qy 1365 CAAATATGTAACCTGTTTGTAAAGAAAGAGCAACATCAATGAAAGAGCTTAAGAAATTC 1424  
 Db 1141 CAAATATGTAACCTGTTTGTAAAGAAAGAGCAACATCAATGAAAGAGCTTAAGAAATTC 1200  
 Qy 1425 TTGACTCCTCTGCACGCTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484



|||||  
Db 1201 TTGACTCTCTGCAGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1260  
|||||  
QY 1485 AAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544  
|||||  
Db 1261 AAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320  
|||||  
QY 1545 GCATATGTGGTGCATCTACAAACCTGCCGCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604  
|||||  
Db 1321 GCATATGTGGTGCATCTACAAACCTGCCGCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380  
|||||  
QY 1605 ATTATATCCCTTCAGGCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTC 1664  
|||||  
Db 1381 ATTATATCCCTTCAGGCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTC 1440  
|||||  
QY 1665 CTCCTAAGAGGATCTCATTTAGGTAAATTCAGAGGAGACAGACAATGCTGGAAGCTGCA 1724  
|||||  
Db 1441 CTCCTAAGAGGATCTCATTTAGGTAAATTCAGAGGAGACAGACAATGCTGGAAGCTGCA 1500  
|||||  
QY 1725 AAGGCTGGAGATGTGAAACTGTAAATAACTGTGTACTGTTTCAGAGTGTCAACTGCAGA 1784  
|||||  
Db 1501 AAGGCTGGAGATGTGAAACTGTAAATAACTGTGTACTGTTTCAGAGTGTCAACTGCAGA 1560  
|||||  
QY 1785 GACATTTGAAGGCGTCAGTCTACACACTTCATTTTGAGCTGGGTATTAACAGAGTGTCC 1844  
|||||  
Db 1561 GACATTTGAAGGCGTCAGTCTACACACTTCATTTTGAGCTGGGTATTAACAGAGTGTCC 1620  
|||||  
QY 1845 GTGGTGGAAATATCTGTACAGCATGAGCTGATGTGCATGCTAAAGATAAAGAGGAGCCCTT 1904  
|||||  
Db 1621 GTGGTGGAAATATCTGTACAGCATGAGCTGATGTGCATGCTAAAGATAAAGAGGAGCCCTT 1680  
|||||  
QY 1905 GTACCTTTGCACAAATGCATGTTCTATGACAAATATGAAAGTTGCAGAACTCTTGTATAA 1964  
|||||  
Db 1681 GTACCTTTGCACAAATGCATGTTCTTATGACAAATATGAAAGTTGCAGAACTCTTGTATAA 1740  
|||||  
QY 1965 CATGAGCAGTAGTAAATGCTAGCTGATTTATGAAATTTACACCTTTACATCAACAGCA 2024  
|||||  
Db 1741 CATGAGCAGTAGTAAATGCTAGCTGATTTATGAAATTTACACCTTTACATCAACAGCA 1800  
|||||  
QY 2025 GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCAA 2084  
|||||  
Db 1801 GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCAA 1860  
|||||  
QY 2085 AAAACAGGGATGGAATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAA 2144  
|||||  
Db 1861 AAAACAGGGATGGAATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTCAA 1920  
|||||  
QY 2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTPAGATGCTGCCAAGAAGGTTGTTTAGCCAGA 2204  
|||||  
Db 1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTPAGATGCTGCCAAGAAGGTTGTTTAGCCAGA 1980  
|||||  
QY 2205 GTGAAGAAGTTGCTTCTCTGATATGTAATTTGCCGGATACCCAGGCGACAGATTCA 2264  
|||||  
Db 1981 GTGAAGAAGTTGCTTCTCTGATATGTAATTTGCCGGATACCCAGGCGACAGATTCA 2040  
|||||  
QY 2265 ACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAAAGTTGCAGAGTATTGTTACAA 2324  
|||||  
Db 2041 ACACCTTTACATTTAGCAGCTGGTTATATAATTTAGAAAGTTGCAGAGTATTGTTACAA 2100  
|||||  
QY 2325 CACGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACTTATTCCCTTACATAATGCAGCA 2384  
|||||  
Db 2101 CACGAGCTGATGTGAATGCCCAAGCAAAAGGAGGACTTATTCCCTTACATAATGCAGCA 2160  
|||||  
QY 2385 TCTTACGGGCTGATGATGTAGACGCTTACTTAATAAGTATATAAGTATGTCATGTC 2444  
|||||  
Db 2161 TCTTACGGGCTGATGATGTAGACGCTTACTTAATAAGTATATAAGTATGTCATGTC 2220  
|||||  
QY 2445 ACGGCAAAATGGGCTTTACACCTTTTGACGAGAGGCCCAAAAGGAGGAGCAACAGCTT 2504  
|||||  
Db 2221 ACGGCAAAATGGGCTTTACACCTTTTGACGAGAGGCCCAAAAGGAGGAGCAACAGCTT 2280  
|||||  
QY 2505 TGTGCTTTTGTCTAGCCCATGGAGCTACCCGACTCTTTAAAAATCAGGAAGCAAAACA 2564  
|||||

Db 2281 TGTGCTTTTGTCTAGCCCATGGAGCTGACCGACTCTTTAAAAATCAGGAAGGACAAACA 2340  
|||||  
QY 2565 CCTTTAGATTTAGTTTTCAGCGGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCA 2624  
|||||  
Db 2341 CCTTTAGATTTAGTTTTCAGCAGATGATGTCAGGCTCTTCTGACAGCAGCCATGCCCCA 2400  
|||||  
QY 2625 TCTGCTCTGCCCTCTTGTTTACAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGGAGCC 2684  
|||||  
Db 2401 TCTGCTCTGCCCTCTTGTTTACAGCCTCAAGTGTCTCAATGGTGTGAGAAGCCAGGAGCC 2460  
|||||  
QY 2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGACAGCAGCAGT 2744  
|||||  
Db 2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGACAGCAGCAGT 2520  
|||||  
QY 2745 CTTGCAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTACTGTTAGTTCAGTGAACA 2804  
|||||  
Db 2521 CTTGCAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTACTGTTAGTTCAGTGAACA 2580  
|||||  
QY 2805 GAGGTGCTTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864  
|||||  
Db 2581 GAGGTGCTTCCAGTTTGGAGAAAAGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640  
|||||  
QY 2865 TTTCGTAAGGAATCTTGGACTTCAGCACCTTAATGGATATATTTGAGAGAGAAACAGATCACT 2924  
|||||  
Db 2641 TTTCGTAAGGAATCTTGGACTTCAGCACCTTAATGGATATATTTGAGAGAGAAACAGATCACT 2700  
|||||  
QY 2925 TTTCGATGATTTAGTTGAGATGGGCAAGAGAGCTCAAGGAGATTTGGAATCAATGCTTAT 2984  
|||||  
Db 2701 TTTCGATGATTTAGTTGAGATGGGCAAGAGAGCTGAAGGAGATTTGGAATCAATGCTTAT 2760  
|||||  
QY 2985 GGACATAGGCACAAACTAATTAAGGAGTTCGAGAGACTTATCTCCGGCAACAAGGCTT 3044  
|||||  
Db 2761 GGACATAGGCACAAACTAATTAAGGAGTTCGAGAGACTTATCTCCGGCAACAAGGCTT 2820  
|||||  
QY 3045 AACCCATATTTAACTTTGAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCCT 3104  
|||||  
Db 2821 AACCCATATTTAACTTTGAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCCT 2880  
|||||  
QY 3105 GATGATAAAGAGTTTCACTCTGCTGGAGGAGAGATGCAAGTACAGTTCGAGAGACAGA 3164  
|||||  
Db 2881 GATGATAAAGAGTTTCACTCTGCTGGAGGAGAGATGCAAGTACAGTTCGAGAGACAGA 2940  
|||||  
QY 3165 GATGAGGTTCATGCAGGTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAGGTT 3224  
|||||  
Db 2941 GATGAGGTTCATGCAGGTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAGGTT 3000  
|||||  
QY 3225 TGTAAACAAGAACTATGGCAAGATACACTCACCGGAGAAAAGAACTTTCTGAAGAAAAC 3284  
|||||  
Db 3001 TGTAAACAAGAACTATGGCAAGATACACTCACCGGAGAAAAGAACTTTCTGAAGAAAAC 3060  
|||||  
QY 3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTATC 3344  
|||||  
Db 3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCTCTTTGTGAATGCAATTATC 3120  
|||||  
QY 3345 CACAAGGCTTTGATGAAGGATCGGTACATAGGTGGTATGTTTGGAGCTGCGCATTTAT 3404  
|||||  
Db 3121 CACAAGGCTTTGATGAAGGATCGGTACATAGGTGGTATGTTTGGAGCTGCGCATTTAT 3180  
|||||  
QY 3405 TTTGCTGAAAACCTTTCCAAAAGCAATCAATATGATATGGAATTCGAGAGGACTACTGGG 3464  
|||||  
Db 3181 TTTGCTGAAAACCTTTCCAAAAGCAATCAATATGATATGGAATTCGAGAGGACTACTGGG 3240  
|||||  
QY 3465 TGTCCAGTTTCAAAAAGACAGATCTTTTACATTTGCCACAGGAGCTGCTCTTTTGGCGG 3524  
|||||  
Db 3241 TGTCCAGTTTCAAAAAGACAGATCTTTTACATTTGCCACAGGAGCTGCTCTTTTGGCGG 3300  
|||||  
QY 3525 GTAACTTTGGGAAAAGTCTTCTCTGAGTTTCAGTGAATGAAATGGCAATTTCTCTCTCA 3584  
|||||  
Db 3301 GTAACTTTGGGAAAAGTCTTCTCTGAGTTTCAGTGAATGAAATGGCAATTTCTCTCTCA 3360  
|||||  
QY 3585 GGTCTACTCAGTCACTGCTAGTGGCCAGTGTAAATGGCTAGCATTAGCTGAATATGTT 3644  
|||||  
Db 3361 GGTCTACTCAGTCACTGCTAGTGGCCAGTGTAAATGGCTAGCATTAGCTGAATATGTT 3420





QY 1415 TAAAGAAATCTTGACTCCTCTGACAGTGGCATCTGAGAAAGCTCATAAATGATGTTGTGA 1474  
Db 1201 TAAAGAAATCTTGACTCCTCTGACAGTGGCATCTGAGAAAGCTCATAAATGATGTTGTGA 1260  
QY 1475 AGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAGACTCTCT 1534  
Db 1261 AGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAGACTCTCT 1320  
QY 1535 ACACAGAGCTGCATATATGCTGCTACAAACCTGCCGCTACTCCTGAGCTATGGGTG 1594  
Db 1321 ACACAGAGCTGCATATATGCTGCTACAAACCTGCCGCTACTCCTGAGCTATGGGTG 1380  
QY 1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAATGAAAATGT 1654  
Db 1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAATGAAAATGT 1440  
QY 1655 ACAGCAACTCCTCCAGAGGGTATCTCAATTAGGTAAATTCAGAGCGACAGACAATGTCT 1714  
Db 1441 ACAGCAACTCCTCCAGAGGGTATCTCAATTAGGTAAATTCAGAGCGACAGACAATGTCT 1500  
QY 1715 GGAAGCTGCAAGGCTGGAGATGCTGAAACTGTAAAAAACTGTACTGTTTCAGAGTGT 1774  
Db 1501 GGAAGCTGCAAGGCTGGAGATGCTGAAACTGTAAAAAACTGTACTGTTTCAGAGTGT 1560  
QY 1775 CAAGTGCAGAGACATTGAAGGGCGTCAGTCTACACCACCTTCATTTGCGAGCTGGGTATA 1834  
Db 1561 CAAGTGCAGAGACATTGAAGGGCGTCAGTCTACACCACCTTCATTTGCGAGCTGGGTATA 1620  
QY 1835 CAGAGTCCGCTGGTGAATATCTGCTACAGCATGAGCTGATGCTGATGCTAAAGATAA 1894  
Db 1621 CAGAGTCCGCTGGTGAATATCTGCTACAGCATGAGCTGATGCTGATGCTAAAGATAA 1680  
QY 1895 AGAGGCTTGTACCTTTGACAAATGCATGTTCTTATGACATATTAAGTTGCGAAGT 1954  
Db 1681 AGAGGCTTGTACCTTTGACAAATGCATGTTCTTATGACATATTAAGTTGCGAAGT 1740  
QY 1955 TCTTGTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGAGAAATTTACACCTTTTACA 2014  
Db 1741 TCTTGTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGAGAAATTTACACCTTTTACA 1800  
QY 2015 TGAAGCAGCAGCAAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTCGAGA 2074  
Db 1801 TGAAGCAGCAGCAAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTCGAGA 1860  
QY 2075 CCTACCAAAAAACACAGGATGGAATACTCCTTTGGATCTTTGTTAAAGATGGAGATAC 2134  
Db 1861 CCTACCAAAAAACACAGGATGGAATACTCCTTTGGATCTTTGTTAAAGATGGAGATAC 1920  
QY 2135 AGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTTAGATGCTGCCAAGAGGGTTG 2194  
Db 1921 AGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTTAGATGCTGCCAAGAGGGTTG 1980  
QY 2195 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATTTGCGCGGATACCCCAAGG 2254  
Db 1981 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATTTGCGCGGATACCCCAAGG 2040  
QY 2255 CAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCCAGATGA 2314  
Db 2041 CAGACATTCACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCCAGATGA 2100  
QY 2315 TTTTGTACAAACAGGAGCTGATGTAATGCCCAAGCAAGAGGAGCTTTATTCCTTTTACA 2374  
Db 2101 TTTTGTACAAACAGGAGCTGATGTAATGCCCAAGCAAGAGGAGCTTTATTCCTTTTACA 2160  
QY 2375 TAATGAGCATCTTACGGGATGATAGATGAGCAGCTCTACTAATAAAGTATAATGTCATG 2434  
Db 2161 TAATGAGCATCTTACGGGATGATAGATGAGCAGCTCTACTAATAAAGTATAATGTCATG 2220  
QY 2435 TGTCAATGCCCGGCAAAATGGGCTTTACACCTTTTGACAGAGCAGGCCCAAAAGGAGG 2494  
Db 2221 TGTCAATGCCCGGCAAAATGGGCTTTACACCTTTTGACAGAGCAGGCCCAAAAGGAGG 2280

QY 2495 AACACAGCTTTGCTGCTTGTGCTAGCCCTGGAGCTGACCCGACTCTTAAAAATCAGGA 2554  
Db 2281 AACACAGCTTTGCTGCTTGTGCTAGCCCTGGAGCTGACCCGACTCTTAAAAATCAGGA 2340  
QY 2555 AGGACAAACACCTTTAGATTTAGTTTTCAGCGGATGATGCTCAGCGCTCTTCTGACAGCAGC 2614  
Db 2341 AGGACAAACACCTTTAGATTTAGTTTTCAGCGAGATGATGCTCAGCGCTCTTCTGACAGCAGC 2400  
QY 2615 CATGCCCCCACTGCTCTGCCCTCTTGTACAGCCCTCAAGTGTCTCAATGGTGTGAGAAG 2674  
Db 2401 CATGCCCCCACTGCTCTGCCCTCTTGTACAGCCCTCAAGTGTCTCAATGGTGTGAGAAG 2460  
QY 2675 CCAGAGGCCACTGCAGATGCTCTCTCTCAGGTCCATCTAGCCCATCAGCCCTTTCTGC 2734  
Db 2461 CCAGAGGCCACTGCAGATGCTCTCTCTCAGGTCCATCTAGCCCATCAGCCCTTTCTGC 2520  
QY 2735 AGCCAGCAGCTTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCACTAGTTAGTTTC 2794  
Db 2521 AGCCAGCAGCTTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCACTAGTTAGTTTC 2580  
QY 2795 AAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCAGGAGTAGATTTTAG 2854  
Db 2581 AAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCAGGAGTAGATTTTAG 2640  
QY 2855 CATAACTCAATTCCTAAGGAATCTTGGACTTGAGCACCCTAATGGATATATTGAGAGAGA 2914  
Db 2641 CATAACTCAATTCCTAAGGAATCTTGGACTTGAGCACCCTAATGGATATATTGAGAGAGA 2700  
QY 2915 ACAGATCACTTTGGATGATATTAGTTGAGATGGGCAACAAGGAGCTGAAGGAGATTGGAAT 2974  
Db 2701 ACAGATCACTTTGGATGATATTAGTTGAGATGGGCAACAAGGAGCTGAAGGAGATTGGAAT 2760  
QY 2975 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGAGCA 3034  
Db 2761 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGAGCA 2820  
QY 3035 ACAAGTCTTAAACCCATATTAACTTTGAACACCTCTGCTAGTGGAACTTCTTTATAGA 3094  
Db 2821 ACAAGTCTTAAACCCATATTAACTTTGAACACCTCTGCTAGTGGAACTTCTTTATAGA 2880  
QY 3095 TCTGCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTGC 3154  
Db 2881 TCTGCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTGC 2940  
QY 3155 AGAGCACAGAGATGGAGTTCATGAGTGGAACTCTTCAACAGATACAAATTTCTCAAGAT 3214  
Db 2941 AGAGCACAGAGATGGAGTTCATGAGTGGAACTCTTCAACAGATACAAATTTCTCAAGAT 3000  
QY 3215 TCAGAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGGAGAAAAGAGTTTC 3274  
Db 3001 TCAGAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGGAGAAAAGAGTTTC 3060  
QY 3275 TGAAGAAAACCAACCATGCCAATGAACGATGCTTATTCATGGCTCTCCTTTTGTGAA 3334  
Db 3061 TGAAGAAAACCAACCATGCCAATGAACGATGCTTATTCATGGCTCTCCTTTTGTGAA 3120  
QY 3335 TGCAATTTATCCACAAGGCTTTGATGAAAGGATGCTACATAGGTGGTATGTTTGGAGC 3394  
Db 3121 TGCAATTTATCCACAAGGCTTTGATGAAAGGATGCTACATAGGTGGTATGTTTGGAGC 3180  
QY 3395 TGGCAATTTATTTGCTGAAACTCTTCCAAAAGCAATCAATATGATATGGAATGCGAGG 3454  
Db 3181 TGGCAATTTATTTGCTGAAACTCTTCCAAAAGCAATCAATATGATATGGAATGCGAGG 3240  
QY 3455 AGGTACTGGGTCTGCAGTTTCACAAAGACAGATCTTGTACATTTGCCACAGCAGCTGCT 3514  
Db 3241 AGGTACTGGGTCTGCAGTTTCACAAAGACAGATCTTGTACATTTGCCACAGCAGCTGCT 3300  
QY 3515 CTTTTCGCGGTAACTTTGGGAAAGTCTTCTTCTGCAAGTTCAGTGAATGAAATGSCACA 3574  
Db 3301 CTTTTCGCGGTAACTTTGGGAAAGTCTTCTTCTGCAAGTTCAGTGAATGAAATGSCACA 3360  
QY 3575 TTTCTCTCCAGGTCATCACTCACTGCTAGCTAGGCGCCAGCTGTAATGGCTTAGCATTAGC 3634







QY 1065 CTTTCATGAGCCAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAGTTATGTT 1124  
|||||  
Db 841 CTTTCATGAGCCAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTCTCTCTTAAGTTATGTT 900  
QY 1125 GCAGACCCCAACACTGCTCAATTTGTCACAATAAAAGTGCTATAGACTTGCTGCCACACCA 1184  
|||||  
Db 901 GCAGACCCCAACACTGCTCAATTTGTCACAATAAAAGTGCTATAGACTTGCTGCCACACCA 960  
QY 1185 CAGTTAAAAAGAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACCA 1244  
|||||  
Db 961 CAGTTAAAAAGAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACCA 1020  
QY 1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCAT 1304  
|||||  
Db 1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCGAAATGGTGAATTTCAAGCAT 1080  
QY 1305 CCTCAACACATGAACACAGCATGCAATGCTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364  
|||||  
Db 1081 CCTCAACACATGAACACAGCATGCAATGCTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140  
QY 1365 CAAATATGTGAACCTGTGCTAAGAAAGAGCAAAACATCAATGAAAGACTAAAGAAATTC 1424  
|||||  
Db 1141 CAAATATGTGAACCTGTGCTAAGAAAGAGCAAAACATCAATGAAAGACTAAAGAAATTC 1200  
QY 1425 TTGACTCCTCTGCAGCTGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTAGTGGTG 1484  
|||||  
Db 1201 TTGACTCCTCTGCAGCTGGCATCTGAGAAAGCTCATATGATGATRTTGTGAAAGTAGTGGTG 1260  
QY 1485 AAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTTGGTCAGACTTCTCTACACAGAGCT 1544  
|||||  
Db 1261 AAACATGAAGCAAAAGGTTAATGCTCTGGATAATCTTTGGTCAGACTTCTCTACACAGAGCT 1320  
QY 1545 GCATATGTTGGTCATCTACAAACCTGCCCCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604  
|||||  
Db 1321 GCATATGTTGGTCATCTACAAACCTGCCCCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380  
QY 1605 ATTATATCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTC 1664  
|||||  
Db 1381 ATTATATCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTC 1440  
QY 1665 CTCCTAAGAGGTTATCTCATTAGGTAAATTCAGAGGAGACAGACAAATTCGTGGAAGCTGCA 1724  
|||||  
Db 1441 CTCCTAAGAGGTTATCTCATTAGGTAAATTCAGAGGAGACAGACAAATTCGTGGAAGCTGCA 1500  
QY 1725 AAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGA 1784  
|||||  
Db 1501 AAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGCAGA 1560  
QY 1785 GACATTTGAAGGGCGTCAGTCTACACCACCTTCATTTTGACGCTGGGTATACAGAGTGTCC 1844  
|||||  
Db 1561 GACATTTGAAGGGCGTCAGTCTACACCACCTTCATTTTGACGCTGGGTATACAGAGTGTCC 1620  
QY 1845 GTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGSCCTT 1904  
|||||  
Db 1621 GTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTT 1680  
QY 1905 GTACCTTTGCACAATGCATGTTCTTTATGACATATATGAAGTTGCAGAACTTCTTGTTAAA 1964  
|||||  
Db 1681 GTACCTTTGCACAATGCATGTTCTTTATGACATATATGAAGTTGCAGAACTTCTTGTTAAA 1740  
QY 1965 CATGGAGCAGTAGTTAATGCTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCA 2024  
|||||  
Db 1741 CATGGAGCAGTAGTTAATGCTAGCTGATTTATGGAATTTTACACCTTTACATGAAGCAGCA 1800  
QY 2025 GCAAAAGGAAAAATATGAAATTTTGCAAACTTCCTGCTCCAGCATGGTGCACACCCTTACCAA 2084  
|||||  
Db 1801 GCAAAAGGAAAAATATGAAATTTTGCAAACTTCCTGCTCCAGCATGGTGCACACCCTTACCAA 1860  
QY 2085 AAAAAACAGGGATGGAATAACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAA 2144  
|||||  
Db 1861 AAAAAACAGGGATGGAATAACTCCTTTGGATCTGTTTAAAGATGGAGATACAGATATTCAA 1920

QY 2145 GATCTGCTTAGGGGAGATCAGCTTTGCTAGATGCTGCCAAGAAAGGTTGTTTAGCCAGA 2204  
|||||  
Db 1921 GATCTGCTTAGGGGAGATCAGCTTTGCTAGATGCTGCCAAGAAAGAGTTGTTTAGCCAGA 1980  
QY 2205 GTGAAGAAAGTTGCTCTCTCTGATTAATTTGCGCGGATACCCAAAGCAGACATTC 2264  
|||||  
Db 1981 GTGAAGAAAGTTGCTCTCTCTGATTAATTTGCGCGGATACCCAAAGCAGACATTC 2040  
QY 2265 ACACCTTTACATTTAGCAGCTGTTATATAATAATTTAGAAGTTGCGAGATTTGTTTACAA 2324  
|||||  
Db 2041 ACACCTTTACATTTAGCAGCTGTTATATAATAATTTAGAAGTTGCGAGATTTGTTTACAA 2100  
QY 2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTTATTCCTTTTACATAATGCAGCA 2384  
|||||  
Db 2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTTATTCCTTTTACATAATGCAGCA 2160  
QY 2385 TCTTAGGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444  
|||||  
Db 2161 TCTTAGGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220  
QY 2445 ACGGACAAATGGCTTTTCACACCTTTGCACGAGCAGCCCAAAAGGAGGAGACACAGCTT 2504  
|||||  
Db 2221 ACGGACAAATGGCTTTTCACACCTTTGCACGAGCAGCCCAAAAGGAGGAGACACAGCTT 2280  
QY 2505 TGTGCTTTTCTGCTAGCCCATGGAGCTGACCGACTCTTTAAATAATCAGGAAGGACAAACA 2564  
|||||  
Db 2281 TGTGCTTTTCTGCTAGCCCATGGAGCTGACCGACTCTTTAAATAATCAGGAAGGACAAACA 2340  
QY 2565 CTTTAGATTTTAGTTTACAGCGGATGATGTCAGCGCTCTCTGACAGCAGCCATGCCCCCA 2624  
|||||  
Db 2341 CTTTAGATTTTAGTTTACAGCGGATGATGTCAGCGCTCTCTGACAGCAGCCATGCCCCCA 2400  
QY 2625 TCTGCTCTGCCCTCTTGTTCAGAGCCTCAAGTGCTCAATGGTGTGAGAAAGCCAGGAGCC 2684  
|||||  
Db 2401 TCTGCTCTGCCCTCTTGTTCAGAGCCTCAAGTGCTCAATGGTGTGAGAAAGCCAGGAGCC 2460  
QY 2685 ACTGCAGATGCTCTCTCTCAGGTCCTCCTAGCCCATCAAGCCTTCTTCGAGCCAGCAGT 2744  
|||||  
Db 2461 ACTGCAGATGCTCTCTCTCAGGTCCTCCTAGCCCATCAAGCCTTCTTCGAGCCAGCAGT 2520  
QY 2745 CTTTGACAACTTATCTGGGAGTTTTCAGAACTGCTCTTCAGTAGTTAGTTCAAGTGAACA 2804  
|||||  
Db 2521 CTTTGACAACTTATCTGGGAGTTTTCAGAACTGCTCTCAGTAGTTAGTTCAAGTGAACA 2580  
QY 2805 GAGGTTGCTTCCAGTTTGAGAAAAAGAGGTTCCAGGAGTAGATTTTAGCATAAAGTCA 2864  
|||||  
Db 2581 GAGGTTGCTTCCAGTTTGAGAAAAAGAGGTTCCAGGAGTAGATTTTAGCATAAAGTCA 2640  
QY 2865 TTCGTAAGGAATCTTGGACTTGAGCACCCTAATGGATATATTTGAGAGACAGATCAGT 2924  
|||||  
Db 2641 TTCGTAAGGAATCTTGGACTTGAGCACCCTAATGGATATATTTGAGAGACAGATCAGT 2700  
QY 2925 TTTGGATGATTAGTTGAGATGGGCAACAAGGAGCTCAAGGAGATTGGAATCAATGCTTAT 2984  
|||||  
Db 2701 TTTGGATGATTAGTTGAGATGGGCAACAAGGAGCTCAAGGAGATTGGAATCAATGCTTAT 2760  
QY 2985 GGACATAGGCACAAACTAATTAAGAGGTGCGAGAGACTTATCTCCGGACAAAGGCTTT 3044  
|||||  
Db 2761 GGACATAGGCACAAACTAATTAAGAGGTGCGAGAGACTTATCTCCGGACAAAGGCTTT 2820  
QY 3045 AACCCATATTTAACTTTGAACACCTCTGCTAGTGGAAACAAATCTTTAGATCTGCTCCT 3104  
|||||  
Db 2821 AACCCATATTTAACTTTGAACACCTCTGCTAGTGGAAACAAATCTTTAGATCTGCTCCT 2880  
QY 3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGCA 3164  
|||||  
Db 2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGCA 2940  
QY 3165 GATGGAGGTTCATGCGAGGTGGAATCTTTCAACAGATACAAATATTCTCAAGATTTCAGAAAGTT 3224  
|||||  
Db 2941 GATGGAGGTTCATGCGAGGTGGAATCTTTCAACAGATACAAATATTCTCAAGATTTCAGAAAGTT 3000  
QY 3225 TGTAAACAGAAACTATGGGAAAGATACACTACCCGGAGAAAGCAAGCTTCTTGAGAGAAAC 3284





Db 1201 TTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCTCA 1260  
Qy 1683 TTAGTAAATCAGAGGACAGACAAATGCTGGNAGCTGCAAGGCTGGAGATGTCGAA 1742  
Db 1261 TTAGTAAATCAGAGGACAGACAAATGCTGGNAGCTGCAAGGCTGGAGATGTCGAA 1320  
Qy 1743 ACTGTAAAAAAGCTGTGACTGTTTTCAGAGTGTCAACTGCAGAGACATGGAAGGCGTTCAG 1802  
Db 1321 ACTGTAAAAAAGCTGTGACTGTTTTCAGAGTGTCAACTGCAGAGACATGGAAGGCGTTCAG 1380  
Qy 1803 TCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCTA 1862  
Db 1381 TCTACACCACTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGGAAATATCTGCTA 1440  
Qy 1863 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGSCCTTGTACCTTTTCACAAATGCA 1922  
Db 1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGAGGSCCTTGTACCTTTTCACAAATGCA 1500  
Qy 1923 TGTCTTATGACATATTAAGAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAAT 1982  
Db 1501 TGTCTTATGACATATTAAGAGTTGCAGAACTTCTTGTAAACATGGAGCAGTAGTTAAT 1560  
Qy 1983 GTAGCTGATTTATGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAATATGAA 2042  
Db 1561 GTAGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAA 1620  
Qy 2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAAT 2102  
Db 1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAAGAGGATGGAAT 1680  
Qy 2103 ACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGGAGAT 2162  
Db 1681 ACTCCTTTGGATCTTGTAAAGATGGAGATACAGATATTTCAATATCTGCTTAGGGGAGAT 1740  
Qy 2163 GCAGCTTTGTAGATGCTGCCAAGAGGTTGTTTACCCAGAGTGAAGAGTGTCTTCT 2222  
Db 1741 GCAGCTTTGTAGATGCTGCCAAGAGGTTGTTTACCCAGAGTGAAGAGTGTCTTCT 1800  
Qy 2223 CCTGATAATGTAATTCGCCGATACCAAGGAGAGACATTCACACCTTTTACATTTAGCA 2282  
Db 1801 CCTGATAATGTAATTCGCCGATACCAAGGAGAGACATTCACACCTTTTACATTTAGCA 1860  
Qy 2283 GCTGGTTTATTAATTTAGAGTTGCAGAGTATTTGTTTCAACACGAGCTGTATGTAAT 2342  
Db 1861 GCTGGTTTATTAATTTAGAGTTGCAGAGTATTTGTTTCAACACGAGCTGTATGTAAT 1920  
Qy 2343 GCCAAGACAAAGGAGACTTATTCCTTTTACATAATGSCAGCTTACGGGATGTAGAT 2402  
Db 1921 GCCAAGACAAAGGAGACTTATTCCTTTTACATAATGSCAGCTTACGGGATGTAGAT 1980  
Qy 2403 GTAGCAGCTCTACTATAAAGTATAATGTCATGTCTCAATGCCAGGACAAATGGGCTTC 2462  
Db 1981 GTAGCAGCTCTACTATAAAGTATAATGTCATGTCTCAATGCCAGGACAAATGGGCTTC 2040  
Qy 2463 ACACCTTTGCACCAAGCAGCCAAAAGGAGGACAGACAGCTTTGTGCTTTGCTTAGCC 2522  
Db 2041 ACACCTTTGCACCAAGCAGCCAAAAGGAGGACAGACAGCTTTGTGCTTTGCTTAGCC 2100  
Qy 2523 CATGGAGCTGACCGCTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2582  
Db 2101 CATGGAGCTGACCGCTCTTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160  
Qy 2583 GCGGATGATGTACAGGCTCTTTGACAGAGCCATGCCGCCATCTGCTCTGCCCTCTGT 2642  
Db 2161 GCGGATGATGTACAGGCTCTTTGACAGAGCCATGCCGCCATCTGCTCTGCCCTCTGT 2220  
Qy 2643 TACAAGCTCAAGTGTCAATGGTGTGAGAGCCAGGAGCCACTGCAGATGCTCTCT 2702  
Db 2221 TACAAGCTCAAGTGTCAATGGTGTGAGAGCCAGGAGCCACTGCAGATGCTCTCT 2280  
Qy 2703 TCAGGTCCTACTAGCCCTCAAGCCCTTTTGTGACAGCAGCAGCTTGTGACAACTTATCTGGG 2762  
Db 2281 TCAGGTCCTACTAGCCCTCAAGCCCTTTTGTGACAGCAGCAGCTTGTGACAACTTATCTGGG 2340

Qy 2763 AGTTTTTCAGAACTGCTCTTCAGTAGTCTAGTCAAGTGAACAGAGGGTCTTCCAGTTTG 2822  
Db 2341 AGTTTTTCAGAACTGCTCTTCAGTAGTCTAGTCAAGTGAACAGAGGGTCTTCCAGTTTG 2400  
Qy 2823 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTACGATAACTCAATTCGTGAAGAACTTTGGA 2882  
Db 2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTACGATAACTCAATTCGTGAAGAACTTTGGA 2460  
Qy 2883 CTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTTGATGTATTTAGTTGAG 2942  
Db 2461 CTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTTGATGTATTTAGTTGAG 2520  
Qy 2943 ATGGGGCACAGGAGCTGAAGGAGATTTGGAATCAANTGCTTATGGACATAGGACAAACTA 3002  
Db 2521 ATGGGGCACAGGAGCTGAAGGAGATTTGGAATCAANTGCTTATGGACATAGGACAAACTA 2580  
Qy 3003 ATTAAGAGGAGTCGAGAGACTTATCTCCGACAAAGAGTCTTAACCCATATTTAACTTTG 3062  
Db 2581 ATTAAGAGGAGTCGAGAGACTTATCTCCGACAAAGAGTCTTAACCCATATTTAACTTTG 2640  
Qy 3063 AACACCTCTGCTAGTGAACAAATTTTATAGATCTGCTCTCTGATGATAAAGAGTTTCAG 3122  
Db 2641 AACACCTCTGCTAGTGAACAAATTTTATAGATCTGCTCTCTGATGATAAAGAGTTTCAG 2700  
Qy 3123 TCTGTGAGGAGAGATGCAAGGATACAGTTCGAGAGCACAGATGGAGGTCATGCAGGT 3182  
Db 2701 TCTGTGAGGAGAGATGCAAGGATACAGTTCGAGAGCACAGATGGAGGTCATGCAGGT 2760  
Qy 3183 GGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTCGTAACAAAGAACTATGG 3242  
Db 2761 GGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAAGTTCGTAACAAAGAACTATGG 2820  
Qy 3243 GAAAGATACACTACCAGGAGAAAGAGTTCGTAAGAAAAACCAACCACTGCCAATGAA 3302  
Db 2821 GAAAGATACACTACCAGGAGAAAGAGTTCGTAAGAAAAACCAACCACTGCCAATGAA 2880  
Qy 3303 CGAATGCTATTTCAATGCTCTCTTTGTAAGTCAATTTATTTGCTGAAACCTCTTCC 3362  
Db 2881 CGAATGCTATTTCAATGCTCTCTTTGTAAGTCAATTTATTTGCTGAAACCTCTTCC 2940  
Qy 3363 AGGCATGCTACATAGTGTGATGTTTGGAGCTGGCATTATTTTGTGCTGAAACCTCTTCC 3422  
Db 2941 AGGCATGCTACATAGTGTGATGTTTGGAGCTGGCATTATTTTGTGCTGAAACCTCTTCC 3000  
Qy 3423 AAGAGCAATCAATATGATATGAAATTTGGAGGAGTACTGGGTGCTCCAGTTCAAAAGAC 3482  
Db 3001 AAGAGCAATCAATATGATATGAAATTTGGAGGAGTACTGGGTGCTCCAGTTCAAAAGAC 3060  
Qy 3483 AGATCTGTTTACATTTTCCACAGGAGCTGCTCTTTTGGCGGTTAACTTTGGGAAAGTCT 3542  
Db 3061 AGATCTGTTTACATTTTCCACAGGAGCTGCTCTTTTGGCGGTTAACTTTGGGAAAGTCT 3120  
Qy 3543 TTCTCTGAGTTTCAATGCAATGAAATGGCACATTTCTCTCCAGGTCATCACTCAGTCACT 3602  
Db 3121 TTCTCTGAGTTTCAATGCAATGAAATGGCACATTTCTCTCCAGGTCATCACTCAGTCACT 3180  
Qy 3603 GGTAGGCCAGTCTAATGGCTTAGCATTTAGCTGAATATGTTATTTACAGAGGAGACAG 3662  
Db 3181 GGTAGGCCAGTCTAATGGCTTAGCATTTAGCTGAATATGTTATTTACAGAGGAGACAG 3240  
Qy 3663 GCTTATCTCTGAGTATTTAATTTACTTACCAGATTTATGAGGCTGAAGTATGGTCGATGGA 3722  
Db 3241 GCTTATCTCTGAGTATTTAATTTACTTACCAGATTTATGAGGCTGAAGTATGGTCGATGGA 3300  
Qy 3723 TAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGGC 3782  
Db 3301 TAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTGGC 3360  
Qy 3783 CTCAGCTTTTACTCTCTTCTGAAAAAAGAAAAA 3816  
Db 3361 CTCAGCTTTTACTCTCTTCTGAAAAAAGAAAAA 3394

Wed Feb 12 16:08:49 2003

us-09-843-159b-2.rge

Page 33

Search completed: February 11, 2003, 18:26:24  
Job time : 9520.64 secs

---







Db 1621 GCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAATCCTCCAGAGGGTATCT 1680  
Qy 1681 CATTTAGGTAATTCAGAGGCAGACAGCAATTCCTGGAAGCTGCAAGGCTGGAGATGCG 1740  
Db 1681 CATTTAGGTAATTCAGAGGCAGACAGCAATTCCTGGAAGCTGCAAGGCTGGAGATGCG 1740  
Qy 1741 AAACGTGTAATAAACTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAAGGCGTC 1800  
Db 1741 AAACGTGTAATAAACTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAAGGCGTC 1800  
Qy 1801 AGTCTACACCACTTCATTTTTCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGC 1860  
Db 1801 AGTCTACACCACTTCATTTTTCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGC 1860  
Qy 1861 TACAGCATGGAGTGTATGTCATGCTAAAGATAAAGAGGCGCTTGTACTGTTGCACAATG 1920  
Db 1861 TACAGCATGGAGTGTATGTCATGCTAAAGATAAAGAGGCGCTTGTACTGTTGCACAATG 1920  
Qy 1921 CATGTTCTTATGGACATTAATGAAGTTGCGAAGCTTCTTGTAAACATGGAGCAGTAGTTA 1980  
Db 1921 CATGTTCTTATGGACATTAATGAAGTTGCGAAGCTTCTTGTAAACATGGAGCAGTAGTTA 1980  
Qy 1981 ATGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAATATG 2040  
Db 1981 ATGTAGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAATATG 2040  
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGACAGCCCTACCAGGAGGATGGA 2100  
Db 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGACAGCCCTACCAGGAGGATGGA 2100  
Qy 2101 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATTAATCAAGATCTGCTTAGGGAG 2160  
Db 2101 ATACTCCTTTGGATCTTGTAAAGATGGAGATACAGATTAATCAAGATCTGCTTAGGGAG 2160  
Qy 2161 ATGCAGCTTTGTAGATGCTGCCAGAGAGGTTGTTTACCGAGAGTGAAGATGTCCT 2220  
Db 2161 ATGCAGCTTTGTAGATGCTGCCAGAGAGGTTGTTTACCGAGAGTGAAGATGTCCT 2220  
Qy 2221 CTCCTGATTAATGTAATTCGCGCATACCCAGGAGCATTCACACACTTTACATTTAG 2280  
Db 2221 CTCCTGATTAATGTAATTCGCGCATACCCAGGAGCATTCACACACTTTACATTTAG 2280  
Qy 2281 CAGCTGGTTTATTAATAATTTAGAGTTGCGAGATTTTGTTCACACAGGAGCTGATGTA 2340  
Db 2281 CAGCTGGTTTATTAATAATTTAGAGTTGCGAGATTTTGTTCACACAGGAGCTGATGTA 2340  
Qy 2341 ATGCCCCAACAGAGAGGAGCTTATTCCTTTTACATAAATGCAGCATTCCTACGGGATGTA 2400  
Db 2341 ATGCCCCAACAGAGAGGAGCTTATTCCTTTTACATAAATGCAGCATTCCTACGGGATGTA 2400  
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGTCATGTCAATGCCACGACAAATGGGCTT 2460  
Db 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGTCATGTCAATGCCACGACAAATGGGCTT 2460  
Qy 2461 TCACACCTTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520  
Db 2461 TCACACCTTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520  
Qy 2521 CCCATGGAGCTGACCGACTCTTAAATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580  
Db 2521 CCCATGGAGCTGACCGACTCTTAAATTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580  
Qy 2581 CAGCGGATGATGTCAGCGCTCTCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640  
Db 2581 CAGCGGATGATGTCAGCGCTCTCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640  
Qy 2641 GTTACAGGCTCAAGTGTCAATGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700  
Db 2641 GTTACAGGCTCAAGTGTCAATGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700  
Qy 2701 CTTTCAGGCTCACTAGCCCATCAAGCCTTTCTGCAGCAGCAGCTTCTGCAGCAGCAGCTTCTG 2760  
Db 2701 CTTTCAGGCTCACTAGCCCATCAAGCCTTTCTGCAGCAGCAGCTTCTGCAGCAGCAGCTTCTG 2760

Qy 2761 GGAGTTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTT 2820  
Db 2761 GGAGTTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGGAAACAGAGGGTCTTCCAGTT 2820  
Qy 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCGTAAAGAACTTTG 2880  
Db 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTACATAACTCAATTCGTAAAGAACTTTG 2880  
Qy 2881 GACTTGAGCACCTAATCGATATATTTGAGAGAGAGATCACTTTTGATGTATTAGTTG 2940  
Db 2881 GACTTGAGCACCTAATCGATATATTTGAGAGAGAGATCACTTTTGATGTATTAGTTG 2940  
Qy 2941 AGATGGGCGCAACAGAGCTGAAGGAGATTTGGAATCAATTCGTATGACATAGGCACAAAC 3000  
Db 2941 AGATGGGCGCAACAGAGCTGAAGGAGATTTGGAATCAATTCGTATGACATAGGCACAAAC 3000  
Qy 3001 TAATTAAGAGAGTCGAGAGCTTATCTCCGAGCAACAGGCTTTAACCCATATTAACTT 3060  
Db 3001 TAATTAAGAGAGTCGAGAGCTTATCTCCGAGCAACAGGCTTTAACCCATATTAACTT 3060  
Qy 3061 TGAACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCCTGATGATAAAGAGTTTC 3120  
Db 3061 TGAACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCCTGATGATAAAGAGTTTC 3120  
Qy 3121 AGTCTGTGGAGAGAGATGCAAAAGTACAGTTCGAGAGCAGAGATGGAGGTCATGCAG 3180  
Db 3121 AGTCTGTGGAGAGAGATGCAAAAGTACAGTTCGAGAGCAGAGATGGAGGTCATGCAG 3180  
Qy 3181 GTGGAATCTTCAACAGATACAAATTTCTCAAGATTCAGAAGTTTGTAAACAGAAACTAT 3240  
Db 3181 GTGGAATCTTCAACAGATACAAATTTCTCAAGATTCAGAAGTTTGTAAACAGAAACTAT 3240  
Qy 3241 GGGAAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAACCAACACATGCCAATG 3300  
Db 3241 GGGAAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAACCAACACATGCCAATG 3300  
Qy 3301 AACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATTCACAAAGGCTTTTGATG 3360  
Db 3301 AACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATTCACAAAGGCTTTTGATG 3360  
Qy 3361 AAAGCATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAACCTTT 3420  
Db 3361 AAAGCATGCTACATAGTGGTATGTTTGGAGCTGGCATTTATTTGCTGAAACCTTT 3420  
Qy 3421 CCAAAAGCAATCAATATGATGAAATTCGAGAGCTACTGGGTGCCAGTTCCACAAG 3480  
Db 3421 CCAAAAGCAATCAATATGATGAAATTCGAGAGGAGTACTGGGTGCCAGTTCCACAAG 3480  
Qy 3481 ACAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGCCTGGGTAACTTTGGGAAAGT 3540  
Db 3481 ACAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGCCTGGGTAACTTTGGGAAAGT 3540  
Qy 3541 CTTTCTCGAGCTTCAGTGCAATGAAATGGCACATTTCTCTCCAGGTCATCAGTCA 3600  
Db 3541 CTTTCTCGAGCTTCAGTGCAATGAAATGGCACATTTCTCTCCAGGTCATCAGTCA 3600  
Qy 3601 CTGTAGCCCGAGTGAATGCGCTAGCATTTAGCTGAATATGTTTATACAGAGGAGAAC 3660  
Db 3601 CTGTAGCCCGAGTGAATGCGCTAGCATTTAGCTGAATATGTTTATACAGAGGAGAAC 3660  
Qy 3661 AGGCTTATCTCTGAGTATTTAATTTACTTACCAGATTTAGGCGCTGAAGGATGTCGATG 3720  
Db 3661 AGGCTTATCTCTGAGTATTTAATTTACTTACCAGATTTAGGCGCTGAAGGATGTCGATG 3720  
Qy 3721 GATAAATAGTATTTTAAAGAACTAAATTCCTACTGAACTTAAATCATCAAGAGCAGAGTG 3780  
Db 3721 GATAAATAGTATTTTAAAGAACTAAATTCCTACTGAACTTAAATCATCAAGAGCAGAGTG 3780  
Qy 3781 GCCTCTAGCTTTTACTCCTTTGCTGAAAAA 3816  
Db 3781 GCCTCTAGCTTTTACTCCTTTGCTGAAAAA 3816







QY	3181	GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAGGTTGTGAACAAGAACTAT	3240
Db	3208	GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAGGTTGTGAACAAGAACTAT	3267
QY	3241	GGGAAAGATACACTCACCGGAGAAAAAGTTTCTGAAAGAAAAACCAACCATGSCCAATG	3300
Db	3268	GGGAAAGATACACTCACCGGAGAAAAAGTTTCTGAAAGAAAAACCAACCATGSCCAATG	3327
QY	3301	AACGAATGCTATTTTCATGGGTCCTCTTTTGTGAAATGCAATATATCCACAAAGGCTTTGATG	3360
Db	3328	AACGAATGCTATTTTCATGGGTCCTCTTTTGTGAAATGCAATATATCCACAAAGGCTTTGATG	3387
QY	3361	AAAGGCATCGGTACATAGTGTGATGTTTGGAGCTGGCATTTATTTTGCCTGAAAACCTCTT	3420
Db	3388	AAAGGCATCGGTACATAGTGTGATGTTTGGAGCTGGCATTTATTTTGCCTGAAAACCTCTT	3447
QY	3421	CCAAAAGCAATCAATATGTATATGGAATGGAGGAGGTACTGGGTGCCAGTTCCACAAG	3480
Db	3448	CCAAAAGCAATCAATATGTATATGGAATGGAGGAGGTACTGGGTGCCAGTTCCACAAG	3507
QY	3481	ACAGATCTTGTTCATTTGCCACAGCGAGCTGCTCTTTTGC CGGGTAACCTTGGGAAAGT	3540
Db	3508	ACAGATCTTGTTCATTTGCCACAGCGAGCTGCTCTTTTGC CGGGTAACCTTGGGAAAGT	3567
QY	3541	CTTTCCTGCAGTTCAGTGCATGAAATGGCACATTCCTCCTCCAGGTCATCACTAGTCA	3600
Db	3568	CTTTCCTGCAGTTCAGTGCATGAAATGGCACATTCCTCCTCCAGGTCATCACTAGTCA	3627
QY	3601	CTGGTAGGCCCCAGTGTAAATGCCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC	3660
Db	3628	CTGGTAGGCCCCAGTGTAAATGCCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC	3687
QY	3661	AGGCTTATCCTGAGTATTTAAATTTACTTACCAGATATGAGGCCCTGAAGGTATGGTCGATG	3720
Db	3688	AGGCTTATCCTGAGTATTTAAATTTACTTACCAGATATGAGGCCCTGAAGGTATGGTCGATG	3747
QY	3721	GATAAATAGTATTTTAAGAAACTAATTCCTAGCTGAACCTAAATCATCAAGCAGCAGTG	3780
Db	3748	GATAAATAGTATTTTAAGAAACTAATTCCTAGCTGAACCTAAATCATCAAGCAGCAGTG	3807
QY	3781	GCCTCTACGTTTTACTCCCTTTCCTGAAAAAAA	3813
Db	3808	GCCTCTACGTTTTACTCCCTTTCCTGAAAAAAA	3840

### RESULT 3

AAC66825  
ID AAC66825 standard; cDNA; 4275 BP.

AAC66825;

DT 27-FEB-2001 (first entry)

Human tankyrase II coding sequence SEQ ID NO: 5.

Human: tankyrase II; telomere length: signal transduction: ss.

XX Homo sapiens.

XX	Key	Location/Qualifiers
FH		

```

CDS
FT 284...3/84
FT /*tag= a

```

```

ET /product=
ET /transl e

```

XX  
DN WC2000061913-21

XX  
DE  
10-00000

	C
	C
	C
	C
	D
	D
	F
	G
	T
XX	D
XX	D

XX

Db 660 CTCCTCTCCATGAAGCTGCAATTAAAGAAAGATTGATGTTTGGCATTTGCTGTACAGC 719  
Qy 661 ATGGAGCTGAGCAACCACTCCGAAATACAGATGGAAGACAGCACTTGATTTAGCAGATC 720  
Db 720 ATGGAGCTGAGCAACCACTCCGAAATACAGATGGAAGACAGCACTTGATTTAGCAGATC 779  
Qy 721 CATCTGCCAAAGCAGTGCTTACTGTGGAATATPAGAAAGATGAAGTCTTGAAGAGTGCCA 780  
Db 780 CATCTGCCAAAGCAGTGCTTACTGTGGAATATPAGAAAGATGAAGTCTTGAAGAGTGCCA 839  
Qy 781 GCAGTGGCAATGAAGAAATATGATGGCTTACTACACACCACTTAATATGTCACCTGCCAG 840  
Db 840 GCAGTGGCAATGAAGAAATATGATGGCTTACTACACACCACTTAATATGTCACCTGCCAG 899  
Qy 841 CAAGTGATGGCAAGAAAGTCAACTCCATTTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900  
Db 900 CAAGTGATGGCAAGAAAGTCAACTCCATTTACATTTGGCAGCAGGATATAACAGAGTAAAGA 959  
Qy 901 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCCTATGAAGATAAAGTGAATCTGG 960  
Db 960 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCCTATGAAGATAAAGTGAATCTGG 1019  
Qy 961 TACCATTACACAATGCTGTTCTTATGGTCATTATGAAGTAACTGAACCTTTGGTCAAGC 1020  
Db 1020 TACCATTACACAATGCTGTTCTTATGGTCATTATGAAGTAACTGAACCTTTGGTCAAGC 1079  
Qy 1021 ATGGTGCCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGCAAGCTT 1080  
Db 1080 ATGGTGCCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGCAAGCTT 1139  
Qy 1081 CTAAGAACAGGTTGAAAGTATGTTCTCTCTCTTAAAGTATTTGGTGCAGCCCAACACTGC 1140  
Db 1140 CTAAGAACAGGTTGAAAGTATGTTCTCTCTCTTAAAGTATTTGGTGCAGCCCAACACTGC 1199  
Qy 1141 TCAATGTGCACAATRAAGTGCTATAGACTGGCTCCACACCAAGTAAAGAAAGAT 1200  
Db 1200 TCAATGTGCACAATRAAGTGCTATAGACTGGCTCCACACCAAGTAAAGAAAGAT 1259  
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAAGCTGATGTTACTC 1260  
Db 1260 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAGAGAAGCTGATGTTACTC 1319  
Qy 1261 GAATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAACACATGAAGA 1320  
Db 1320 GAATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAACACATGAAGA 1379  
Qy 1321 CAGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTTGAACGT 1380  
Db 1380 CAGCATTTGCATTTGCTGCTGCATCTCCATATCCCAAGAAAGCAATATGTTGAACGT 1439  
Qy 1381 TGCTAAGAAAGAGCAACATCAATGAAAGACTTAAAGAAATTCCTGACTCTCTGCAAGC 1440  
Db 1440 TGCTAAGAAAGAGCAACATCAATGAAAGACTTAAAGAAATTCCTGACTCTCTGCAAGC 1499  
Qy 1441 TGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGTGGAACATGAAGCAAGG 1500  
Db 1500 TGGCATCTGAGAAAGCTCATATGATGTTGTTGAAGTGTGGAACATGAAGCAAGG 1559  
Qy 1501 TTAATGCTCTGGATATCTTGGCTGAGACTTCTACACAGAGCTGCATATGTTGGTCAATC 1560  
Db 1560 TTAATGCTCTGGATATCTTGGCTGAGACTTCTACACAGAGCTGCATATGTTGGTCAATC 1619  
Qy 1561 TACAAACCTGCCCTTACTCTCTGAGCTATGGGTGATGCTTAACATATATCCCTTCAGG 1620  
Db 1620 TACAAACCTGCCCTTACTCTCTGAGCTATGGGTGATGCTTAACATATATCCCTTCAGG 1679  
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCCAAAGAGGATCT 1680  
Db 1680 GCTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCTCTCCAAAGAGGATCT 1739  
Qy 1681 CATTAGGTAAATCAGAGGAGACAGACAAATTCCTGGAAGCTCCAAAGGCTGGAGATGCG 1740  
Db 1740 CATTAGGTAAATCAGAGGAGACAGACAAATTCCTGGAAGCTCCAAAGGCTGGAGATGCG 1799

Qy 1741 AAACGTGTAATAAATACTGTGTAAGTCTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGTC 1800  
Db 1800 AAACGTGTAATAAATACTGTGTAAGTCTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGTC 1859  
Qy 1801 AGTCTACACCACTTCAATTTTGCAGCTGGGTATACAGAGTGTCCGGTGGTGAATATCTGC 1860  
Db 1860 AGTCTACACCACTTCAATTTTGCAGCTGGGTATACAGAGTGTCCGGTGGTGAATATCTGC 1919  
Qy 1861 TACAGCATGAGCTGATGTCATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATG 1920  
Db 1920 TACAGCATGAGCTGATGTCATGCTAAAGATAAAGAGGCTTGTACCTTTGCACAATG 1979  
Qy 1921 CATCTCTTATGGACATTTAAGTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTA 1980  
Db 1980 CATCTCTTATGGACATTTAAGTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTA 2039  
Qy 1981 ATGTAGCTGATTTATGAAATTTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATG 2040  
Db 2040 ATGTAGCTGATTTATGAAATTTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATG 2099  
Qy 2041 AAATTTGCAAACTTCTCTCCAGCATGTCAGACCTTACCAAAAACACAGGATGGAA 2100  
Db 2100 AAATTTGCAAACTTCTCTCCAGCATGTCAGACCTTACCAAAAACACAGGATGGAA 2159  
Qy 2101 ATACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAG 2160  
Db 2160 ATACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGAG 2219  
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTAGCCAGAGTGAAGATGTTCTTT 2220  
Db 2220 ATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTAGCCAGAGTGAAGATGTTCTTT 2279  
Qy 2221 CTCTGTATATGTAATTTGCCGATACCAAGCAGACATTCACACCTTTTACATTTAG 2280  
Db 2280 CTCTGTATATGTAATTTGCCGATACCAAGCAGACATTCACACCTTTTACATTTAG 2339  
Qy 2281 CAGCTGGTTATTAATTTAGAGTTTGCAGAGTATTTGTTACACACGAGCTGATGTGA 2340  
Db 2340 CAGCTGGTTATTAATTTAGAGTTTGCAGAGTATTTGTTACACACGAGCTGATGTGA 2399  
Qy 2341 ATGCCCAAGCAAGAGGAGGACTTATCTTTTACATATGAGCATCTTACGGGATGTAG 2400  
Db 2400 ATGCCCAAGCAAGAGGAGGACTTATCTTTTACATATGAGCATCTTACGGGATGTAG 2459  
Qy 2401 ATGTAGCAGCTCTACTTAATAAGTATTAATGATGTGCAATGCCACGACAAATGGGCTT 2460  
Db 2460 ATGTAGCAGCTCTACTTAATAAGTATTAATGATGTGCAATGCCACGACAAATGGGCTT 2519  
Qy 2461 TCACACCTTTTGCAGAGCAGCCCAAGGAGCAGCAACACAGCTTTGTGCTTTGTGCTAG 2520  
Db 2520 TCACACCTTTTGCAGAGCAGCCCAAGGAGCAGCAACACAGCTTTGTGCTTTGTGCTAG 2579  
Qy 2521 CCCATGAGCTGAGCCGACCTCTTAAATAATCAGGAAGCAGCAACACCTTTAGATTTAGTTT 2580  
Db 2580 CCCATGAGCTGAGCCGACCTCTTAAATAATCAGGAAGCAGCAACACCTTTAGATTTAGTTT 2639  
Qy 2581 CAGCGGATGATGTCAGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCGCTCTT 2640  
Db 2640 CAGCGGATGATGTCAGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCGCTCTT 2699  
Qy 2641 GTTACAAGCCTTCAAGTGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCT 2700  
Db 2700 GTTACAAGCCTTCAAGTGTCTCAATGGTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCT 2759  
Qy 2701 CTTCAGGTCCATCTAGGCCATCAAGCCTTTCTGAGCCAGCAGCTCTTGACAACTTATCTG 2760  
Db 2760 CTTCAGGTCCATCTAGGCCATCAAGCCTTTCTGAGCCAGCAGCTCTTGACAACTTATCTG 2819  
Qy 2761 GGAGTTTTTCAGAACTGTCTTCAAGTGTAGTTAGTTCAAGTGGAAACAGAGGCTTCCAGTT 2820  
Db 2820 GGAGTTTTTCAGAACTGTCTTCAAGTGTAGTTAGTTCAAGTGGAAACAGAGGCTTCCAGTT 2879

```
QY 2821 TGAGAAAAAGGAGGTCCAGGAGTAGATTTTASCATAACTCAATTCGTAAAGAACTTTG 2880
Db 2880 TGAGAAAAAGGAGGTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAAGAACTTTG 2939
QY 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATAGTTG 2940
Db 2940 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATAGTTG 2999
QY 2941 AGATGGGACAAAGGAGCTGAAGGAGATGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db 3000 AGATGGGACAAAGGAGCTGAAGGAGATGGAATCAATGCTTATGGACATAGGCACAAAC 3059
QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTTAACCCATATTTAACTT 3060
Db 3060 TAATTAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGCTTTAACCCATATTTAACTT 3119
QY 3061 TGAACACCTCTGTAGTGAACAATTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTC 3120
Db 3120 TGAACACCTCTGTAGTGAACAATTTCTTATAGATCTGTCTCTGATGATAAAGAGTTTC 3179
QY 3121 AGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGAGAGCACAGAGATGAGGTGTCAG 3180
Db 3180 AGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGAGAGCACAGAGATGAGGTGTCAG 3239
QY 3181 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAGTTTCTAAACAAGAACTAT 3240
Db 3240 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAGAAGTTTCTAAACAAGAACTAT 3299
QY 3241 GGAAGAGATACACTCACCGGAGAAAAGATTTCTGAAGAAACCAACCAACCATGCCAATG 3300
Db 3300 GGAAGAGATACACTCACCGGAGAAAAGATTTCTGAAGAAACCAACCAACCATGCCAATG 3359
QY 3301 AACGAATGCTATTTCACTGCTCTCTTTTGTGAATCAATTAATCCACAAGGCTTTGATG 3360
Db 3360 AACGAATGCTATTTCACTGCTCTCTTTTGTGAATCAATTAATCCACAAGGCTTTGATG 3419
QY 3361 AAAGGATCGGTACATAGTGGTATGTTGGAGCTGGCAITATTTTGTGAAACTCTT 3420
Db 3420 AAAGGATCGGTACATAGTGGTATGTTGGAGCTGGCAITATTTTGTGAAACTCTT 3479
QY 3421 CCAAAAGCAATCAATATGTATATGGAATGGAGGAGTACTGGGTCTCCAGTTTCAAAAG 3480
Db 3480 CCAAAAGCAATCAATATGTATATGGAATGGAGGAGTACTGGGTCTCCAGTTTCAAAAG 3539
QY 3481 ACAGATCTTGTACATTTGCCACAGCAGTGTCTTTTGGCGGGTAACCTTTGGGAAAGT 3540
Db 3540 ACAGATCTTGTACATTTGCCACAGCAGTGTCTTTTGGCGGGTAACCTTTGGGAAAGT 3599
QY 3541 CTTTCTGTCAGTTTCAGTGAATGAAATGGCAATTCCTCCAGTCACTCAGTCA 3600
Db 3600 CTTTCTGTCAGTTTCAGTGAATGAAATGGCAATTCCTCCAGTCACTCAGTCA 3659
QY 3601 CTGGTAGGCCCCAGTGAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db 3660 CTGGTAGGCCCCAGTGAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAAC 3719
QY 3661 AGCTTATCCTGAGTATTTAATTTACTTACAGATATAGGCGCTGAAGGTATGGTGCATG 3720
Db 3720 AGCTTATCCTGAGTATTTAATTTACTTACAGATATAGGCGCTGAAGGTATGGTGCATG 3779
QY 3721 GATAAATAGTATTTTAAAGAACTAATTCCTGAACTTCAATCAATCAATCAAGCAGAGTG 3780
Db 3780 GATAAATAGTATTTTAAAGAACTAATTCCTGAACTTCAATCAATCAATCAAGCAGAGTG 3839
QY 3781 GCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3813
Db 3840 GCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3872
```

RESULT 4  
AAF63930  
ID AAF63930 standard; DNA: 4406 BP.  
XX

```
AC AAF63930;
XX 05-APR-2001 (first entry)
DE Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 106.
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX Homo sapiens.
XX WO200100849-A1.
PN 04-JAN-2001.
XX 28-JUN-2000; 2000WO-US17827.
PF 29-JUN-1999; 99US-0141582.
PR (ICOS-) ICOS CORP.
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
PI WPI; 2001-102896/11.
XX P-PSDB; AAB66290.
DR New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX Example 2; Page 168-173; 242pp; English.
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX Sequence 4406 BP; 1268 A; 910 C; 1054 G; 1174 T; 0 other;
SQ
Query Match 99.4%; Score 3791.4; DB 22; Length 4406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1 CGCGCTGCTCCCGCCCGCGGCGAGCGGGGCGAGGAGCCAGGAGGGCGCGCG 60
Db 67 CGCGCTGCTCCCGCCCGCGGCGAGCGGGGCGAGGAGCCAGGAGGGCGCGCG 126
QY 61 TGGGCGCGCGCCATGGGACTGCGCGGATCGCGTGACAGCAGGAGCGCCGCGGG 120
Db 127 TGGGCGCGG-CATGGGACTGCGCGGATCGCGTGACAGCAGGAGCGCCGCGGG 185
QY 121 CCTGAGCGGCTTCTCCGGGGGCGCTCCCGCTCTCTGCTCGCGGGGCGGGCTCTCG 180
Db 186 CCTGAGCGGCTTCTCCGGGGGCGCTCCCGCTCTCTGCTCGCGGGGCGGGCTCTCG 245
QY 181 TCCGGTGTGGCGCTGTTGCTGGCTGTGGCGCGCGCGCGGAGGAGGAGGAGGAG 240
Db 246 TCCGGTGTGGCGCTGTTGCTGGCTGTGGCGCGCGCGGAGGAGGAGGAGGAGGAG 305
QY 241 GCGCGCGCGGGGAGGCGGCTGCGCGAGCGCGCGCGCGGAGGAGGAGGAGGAGGAG 300
Db 306 GCGCGCGCGGGGAGGCGGCTGCGCGAGCGCGCGCGCGGAGGAGGAGGAGGAGGAG 365
QY 301 GAGAGCTGTTCAGGCGTCCCGCAACGGGAGCGTGAAGAGTCAAGAGGCTGGTGAACG 360
Db 366 GAGAGCTGTTCAGGCGTCCCGCAACGGGAGCGTGAAGAGTCAAGAGGCTGGTGAACG 425
QY 361 CTGAGAAGGTGAACAGCGCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
```

Db 426 CTGAGAGGTGAACCGCGCAGACGGGGGCGAGGAAATCCACCCGCTGCATCTCGCG 485  
Qy 421 CAGGTTTGGCGGGAAGACGCTAGTTGAATATTTCTTTCAGAAATGGTCCAAATGTCGAAG 480  
Db 486 CAGGTTTGGCGGGAAGACGCTAGTTGAATATTTCTTTCAGAAATGGTCCAAATGTCGAAG 545  
Qy 481 CACGTGATGATGGGGCCCTTATTCCTCTTCATAATGCAATGCTCTTTTGGTCATGCTGAAG 540  
Db 546 CACGTGATGATGGGGCCCTTATTCCTCTTCATAATGCAATGCTCTTTTGGTCATGCTGAAG 605  
Qy 541 TAGTCAATCTCTTTTGGACATGCTGCAGACCCCAATGCTCGAGATAATTTGGAATTAATA 600  
Db 606 TAGTCAATCTCTTTTGGACATGCTGCAGACCCCAATGCTCGAGATAATTTGGAATTAATA 665  
Qy 601 CTCCTCTCCATCAAGCTGCAATTAAGGAAAGATTGATTTTGGCATTTGCTGTTTACAGC 660  
Db 666 CTCCTCTCCATCAAGCTGCAATTAAGGAAAGATTGATTTTGGCATTTGCTGTTTACAGC 725  
Qy 661 ATGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 720  
Db 726 ATGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGGACAGCATTTGGATTTAGCAGATC 785  
Qy 721 CATCTGCCAAACGAGTCTTACTGTTGAATATAAGAAAGATGAACCTCTTAGAAGTGCCA 780  
Db 786 CATCTGCCAAACGAGTCTTACTGTTGAATATAAGAAAGATGAACCTCTTAGAAGTGCCA 845  
Qy 781 GGAGTGGCAATGAAGAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCACG 840  
Db 846 GGAGTGGCAATGAAGAAATGATGGCTCTACTCACACCATTAATGTCACCTGCCACG 905  
Qy 841 CAAGTGATGGCAAGTCAACTCCATTAACATTTGGCAGCAGGATATAACAGAGTAAAGA 900  
Db 906 CAAGTGATGGCAAGTCAACTCCATTAACATTTGGCAGCAGGATATAACAGAGTAAAGA 965  
Qy 901 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCATGCTTAAAGTAAAGGTGATCTGG 960  
Db 966 TTGTACAGCTGTTACTGCAACATGAGCTGATGTCATGCTTAAAGTAAAGGTGATCTGG 1025  
Qy 961 TACCATTAACAAATGCTCTTCTTATGGTCAATTAAGTAACTGAACCTTTTGGTCAAGC 1020  
Db 1026 TACCATTAACAAATGCTCTTCTTATGGTCAATTAAGTAACTGAACCTTTTGGTCAAGC 1085  
Qy 1021 ATGGTCCCTGTGTAAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGAGCTTT 1080  
Db 1086 ATGGTCCCTGTGTAAATGCAATGGACTTGTGGCAATTCACCTCTTCATGAGGAGCTTT 1145  
Qy 1081 CTAAAGAACAGGTTGAAGTATGTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140  
Db 1146 CTAAAGAACAGGTTGAAGTATGTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1205  
Qy 1141 TCAATTTGTACAAATAAAGTGTATAGACTTGGCTCCCAACACAGTTTAAAGAAAGAT 1200  
Db 1206 TCAATTTGTACAAATAAAGTGTATAGACTTGGCTCCCAACACAGTTTAAAGAAAGAT 1265  
Qy 1201 TAGCATATGAATTTAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTTC 1260  
Db 1266 TAGCATATGAATTTAAGGCCACTCGTTGCTGCAAGCTGCAGAGAGCTGATGTTACTTC 1325  
Qy 1261 GAATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAAAACACATGAA 1320  
Db 1326 GAATCAAAAACATCTCTCTCTGGAATGGTGAATTTCAAGCATCCTCAAAACACATGAA 1385  
Qy 1321 CAGCATTTGATTTGCTGCTGCTATCCATATCCCAAAAAGAACAAATATGTGAAGTCT 1380  
Db 1386 CAGCATTTGATTTGCTGCTGCTATCCATATCCCAAAAAGAACAAATATGTGAAGTCT 1445  
Qy 1381 TGCTAAGAAAAGGAGCAACATCAATGAAGAACTAAAGAAATTCCTTGACTCTCTGCAAG 1440  
Db 1446 TGCTAAGAAAAGGAGCAACATCAATGAAGAACTAAAGAAATTCCTTGACTCTCTGCAAG 1505  
Qy 1441 TGCGATCTGAGAAAGCTCATAATGATGTTGTTGAAGTGTGTTGAACATGAAGCAAGG 1500  
Db 1506 TGCGATCTGAGAAAGCTCATAATGATGTTGTTGAAGTGTGTTGAACATGAAGCAAGG 1565

Qy 1501 TTAATGCTCTGGATAATCTTTGGTCAGACTTCTCTACAGAGCTGCATATTTGGTTCATC 1560  
Db 1566 TTAATGCTCTGGATAATCTTTGGTCAGACTTCTCTACAGAGCTGCATATTTGGTTCATC 1625  
Qy 1561 TACAAACCTCGCGCTACTCCTGAGCTATGGGTGTGATCTTAACATTTATTCCTTCAGG 1620  
Db 1626 TACAAACCTCGCGCTACTCCTGAGCTATGGGTGTGATCTTAACATTTATTCCTTCAGG 1685  
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGTAAGCAACTCTCTCAAGAGGGTATCT 1680  
Db 1686 GCTTTACTGCTTTACAGATGGGAAATGTAAGCAACTCTCTCAAGAGGGTATCT 1745  
Qy 1681 CATTAGGTAAATTCAGAGCCAGACACAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCG 1740  
Db 1746 CATTAGGTAAATTCAGAGCCAGACACAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCG 1805  
Qy 1741 AAACGTAAAAAAACTGTACTGTTTACAGAGTGTCAACTGCAGAGACATTTGAAGGGGCTC 1800  
Db 1806 AAACGTAAAAAAACTGTACTGTTTACAGAGTGTCAACTGCAGAGACATTTGAAGGGGCTC 1865  
Qy 1801 AGTCTACACCACTTCAATTTTGCAGCTGGGTATAACAGAGTGTCCGTGCTGGAATATCTGC 1860  
Db 1866 AGTCTACACCACTTCAATTTTGCAGCTGGGTATAACAGAGTGTCCGTGCTGGAATATCTGC 1925  
Qy 1861 TACAGCATGGAGCTGATGTGCTAAAGATAAAGAGGCGCTTTGTACCTTTGCACAATG 1920  
Db 1926 TACAGCATGGAGCTGATGTGCTAAAGATAAAGAGGCGCTTTGTACCTTTGCACAATG 1985  
Qy 1921 CATGTTCTTATGGACATTTAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTA 1980  
Db 1986 CATGTTCTTATGGACATTTAAGTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTA 2045  
Qy 1981 ATGTAGCTGATTTATGAAATTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATG 2040  
Db 2046 ATGTAGCTGATTTATGAAATTTACACTTTTACATGAAGCAGCAGCAAAAGGAAATATG 2105  
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCACAAAAAACAAGGATGAA 2100  
Db 2106 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCACAAAAAACAAGGATGAA 2165  
Qy 2101 ATACTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAG 2160  
Db 2166 ATACTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAG 2225  
Qy 2161 ATGCACTTTTCTAGATGCTGCCAAGAGGCTTGTGTAGCCAGAGTGAAGAGTTGCTTT 2220  
Db 2226 ATGCACTTTTCTAGATGCTGCCAAGAGGCTTGTGTAGCCAGAGTGAAGAGTTGCTTT 2285  
Qy 2221 CTCTGTAAATGTAAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAG 2280  
Db 2286 CTCTGTAAATGTAAATTTGCCGCGATACCCAAAGCAGACATTTCAACACCTTTACATTTAG 2345  
Qy 2281 CAGCTGGTTTAAATTAATTTAGAAATTTAGAAATTTAGAAATTTAGAAATTTAGAAATTTAG 2340  
Db 2346 CAGCTGGTTTAAATTAATTTAGAAATTTAGAAATTTAGAAATTTAGAAATTTAGAAATTTAG 2405  
Qy 2341 ATGCCAAGACAAAGAGGAGCTTTATTCCTTTTACATAATGCAAGTCTTTACGGGCTAGTAG 2400  
Db 2406 ATGCCAAGACAAAGAGGAGCTTTATTCCTTTTACATAATGCAAGTCTTTACGGGCTAGTAG 2465  
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCATGTCACCGGACAAATGGGCTT 2460  
Db 2466 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTGTCATGTCACCGGACAAATGGGCTT 2525  
Qy 2461 TCACACCTTTGCAGAACAGCCCAAGAGGACGACACACCTTTTGCTTTTGCTAGTAG 2520  
Db 2526 TCACACCTTTGCAGAACAGCCCAAGAGGACGACACACCTTTTGCTTTTGCTAGTAG 2585  
Qy 2521 CCCATGGAGCTGACCCGACTCTTTAAATATCAGGAAGGACGACACACCTTTAGATTTAGTTT 2580  
Db 2586 CCCATGGAGCTGACCCGACTCTTTAAATATCAGGAAGGACGACACACCTTTAGATTTAGTTT 2645

QY 2581 CAGCGGATGATGTGAGCGCTCTTCTGTACAGAGCCATGCCCCCATCTGCTGCTGCTCTT 2640  
Db 2646 CAGCAGATGATGTGAGCGCTCTTCTGACAGAGCCATGCCCCCATCTGCTGCTGCTCTT 2705  
QY 2641 GTTACAAGCCTCAAGTCTCAATGGTGTGAGAGCCCGAGAGCCACTGCGAGATGCTCTCT 2700  
Db 2706 GTTACAAGCCTCAAGTCTCAATGGTGTGAGAGCCCGAGAGCCACTGCGAGATGCTCTCT 2765  
QY 2701 CTTTCAGTCCATCTAGCCCATCAAGCCCTTCTGTCAGCCAGCAGCTTGTGACAACCTTATCTG 2760  
Db 2766 CTTTCAGTCCATCTAGCCCATCAAGCCCTTCTGTCAGCCAGCAGCTTGTGACAACCTTATCTG 2825  
QY 2761 GGAGTTTTTTCAGACTGCTCTTCAGTAGTATAGTTCAAGTGAAGAGAGGGTCTTCCAGTT 2820  
Db 2826 GGAGTTTTTTCAGACTGCTCTTCAGTAGTATAGTTCAAGTGAAGAGAGGGTCTTCCAGTT 2885  
QY 2821 TGGAGAAAAGAGGGTTCAGAGGAGTAGATTTAGCATTAACCTCAATTCGTAAAGNAICTTG 2880  
Db 2886 TGGAGAAAAGAGGGTTCAGAGGAGTAGATTTAGCATTAACCTCAATTCGTAAAGNAICTTG 2945  
QY 2881 GACTTGGACACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATAGTTG 2940  
Db 2946 GACTTGGACACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATAGTTG 3005  
QY 2941 AGATGGGCACAGGAGCTGAAGAGATGGAATCAATGCTTATGGACATAGGCACAAAC 3000  
Db 3006 AGATGGGCACAGGAGCTGAAGAGATGGAATCAATGCTTATGGACATAGGCACAAAC 3065  
QY 3001 TAAATTAAGGAGTCGAGAGACTTATCTCCGACAAAGCTCTTAACCCATATTTAACTT 3060  
Db 3066 TAAATTAAGGAGTCGAGAGACTTATCTCCGACAAAGCTCTTAACCCATATTTAACTT 3125  
QY 3061 TGAACACCTCTGCTAGTGGAAACATTTATAGATCTGCTCCTGATGATAAAGAGTTTC 3120  
Db 3126 TGAACACCTCTGCTAGTGGAAACATTTATAGATCTGCTCCTGATGATAAAGAGTTTC 3185  
QY 3121 AGTCTGTGGAGGAGAGATGCAAGTACAGTTCGAGAGCACAGATGGAGTTCATGCAG 3180  
Db 3186 AGTCTGTGGAGGAGAGATGCAAGTACAGTTCGAGAGCACAGATGGAGTTCATGCAG 3245  
QY 3181 GTGGAATCTTCAACAGATACAAATATTTCTCAAGATTCTCAGAGGTTTGTACAAGAACTAT 3240  
Db 3246 GTGGAATCTTCAACAGATACAAATATTTCTCAAGATTCTCAGAGGTTTGTACAAGAACTAT 3305  
QY 3241 GGAAGAGATACACTCCCGGAGAAAAAGTTTCTGAAGAAAAACCAACCATGCCAATG 3300  
Db 3306 GGAAGAGATACACTCCCGGAGAAAAAGTTTCTGAAGAAAAACCAACCATGCCAATG 3365  
QY 3301 AACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCACAAGGCTTTGATG 3360  
Db 3366 AACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCACAAGGCTTTGATG 3425  
QY 3361 AAAGGATGCTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGTGAAACTCTT 3420  
Db 3426 AAAGGATGCTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGTGAAACTCTT 3485  
QY 3421 CCAAAAGCAATCAATATGATATATGGAATTTGGAGAGGTACTGGGTGTCAGATTTCACAAAG 3480  
Db 3486 CCAAAAGCAATCAATATGATATATGGAATTTGGAGAGGTACTGGGTGTCAGATTTCACAAAG 3545  
QY 3481 ACAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACTTGGGAAAGT 3540  
Db 3546 ACAGATCTTTGTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACTTGGGAAAGT 3605  
QY 3541 CTTTCTGTCAGTTTCAGTGCATGAAATGGACATTTCTCTCCAGTGCATCACTCAGTCA 3600  
Db 3606 CTTTCTGTCAGTTTCAGTGCATGAAATGGACATTTCTCTCCAGTGCATCACTCAGTCA 3665  
QY 3601 CTGGTAGGCCCGAGTGPAAATGGCCTAGCATATAGCTGAATATGTTATTTACAGAGAGAAC 3660  
Db 3666 CTGGTAGGCCCGAGTGPAAATGGCCTAGCATATAGCTGAATATGTTATTTACAGAGAGAAC 3725  
QY 3661 AGGCTTATCTTGAATATTTAACTTACTTACCAGATTAAGGCCCTGAAGTATGGTCGATG 3720

Db 3726 AGGCTTATCTCTGAGTATTTAACTTACCAGATTATGAGCCCTGAAGGTATGTCGATG 3785  
QY 3721 GATAAATAGTATTTTAAAGAACTAATTCCTACTGAACCTAAATCATCAAGCAGCAGTG 3780  
Db 3786 GATAAATAGTATTTTAAAGAACTAATTCCTACTGAACCTAAATCATCAAGCAGCAGTG 3845  
QY 3781 GCCTCTACGTTTACTCTCTTGTGCTGAAAAAAA 3813  
Db 3846 GCCTCTACGTTTACTCTCTTGTGCTGAAAAAAA 3878

RESULT 5

AAF63953

ID AAF63953 standard; DNA; 4992 BP.

XX AAF63953;

AC

DT 05-APR-2001 (first entry)

XX Human tankyrase2 TANK2-SHORT coding sequence SEQ ID NO: 134.

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder; ds.

XX Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-0517827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

WPI: 2001-102896/11.

P-PSDB; AAB66295.

New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders.

Claim 6; Page 195-200; 242pp; English.

The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

Sequence 4992 BP; 1454 A; 1064 C; 1191 G; 1283 T; 0 other;

XX

SQ

Query Match

Best Local Similarity 99.4%; Score 3791.4; DB 22; Length 4992;

Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 CGCGCTGCTCCCGCCCGCGGCGAGCGGGGCGAGGCCAGGAGGCCGCGCG 60

Db 653 CGCGCTGCTCCCGCCCGCGGCGAGCGGGGCGAGGCCAGGAGGCCGCGCG 712

QY 61 TGGGCGCGCCCATGCGACTGCGCGGATCCGGTGCAGCAGGAGGCCAAGCGCCCGG 120

Db 713 TGGGCGCGG-CCATGGACTGCGCGGATCCGGTGCAGCAGGAGGCCAAGCGCCCGG 771

QY 121 CCTGAGCGCGTCTTCGCGGGGGGCTCGCCCTCTCTGTCGCGGGGCGGCTCTCTG 180

Db 772 CCTGTAGCGCGTCTTCTCCGGGGGCGCTCGCCCTCTCTGTCTCGGGGCGCGGCGCTCCTGC 831  
QY 181 TCCGGTGTGTGGCGCTGTTCCTGGCTGTGCGCGCGCCAGGATCATGTGCGGTCCGCCGT 240  
Db 832 TCCGGTGTGTGGCGCTGTTCCTGGCTGTGCGCGCGCCAGGATCATGTGCGGTCCGCCGT 891  
QY 241 GCCCCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
Db 892 GCCCCGCGGGGAGCGGCTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 951  
QY 301 GAGAGCTGTTTCGAGGCGTGC CGCAACGGGAGCGTGAAGCAGTCAAGAGCGTGTGAAGC 360  
Db 952 GAGAGCTGTTTCGAGGCGTGC CGCAACGGGAGCGTGAAGCAGTCAAGAGCGTGTGAAGC 1011  
QY 361 CTGAGAAGGTGAACGCCGCGACACCGCGCGGAGGAAATCCACCGCGTGCATCTTCGCG 420  
Db 1012 CTGAGAAGGTGAACGCCGCGACACCGCGCGGAGGAAATCCACCGCGTGCATCTTCGCG 1071  
QY 421 CAGGTTTTGGCGGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAATGTCCAAG 480  
Db 1072 CAGGTTTTGGCGGGAAGACGTAGTTGAATATTTGCTTCAGAAATGGTGCAAATGTCCAAG 1131  
QY 481 CAGCTGATGATGGGCGCTTATTCCTTCATTAAGTCATGCTCTTTTGGTCAATGCTGAAG 540  
Db 1132 CAGCTGATGATGGGCGCTTATTCCTTCATTAAGTCATGCTCTTTTGGTCAATGCTGAAG 1191  
QY 541 TACTCAATCTCCTTTTGCACATGTTGCGACCGCGAGCCCAATGCTCGAGATAATTCGAATATA 600  
Db 1192 TAGTCAATCTCCTTTTGGCGACATGTTGCGACCGCCCAATGCTCGAGATAATTCGAATATA 1251  
QY 601 CTCCTCTCCATGAAGTGC AATTAAGGAAAGATTGATGTTTGCATTTGTGCTTACAGC 660  
Db 1252 CTCCTCTCCATGAAGTGC AATTAAGGAAAGATTGATGTTTGCATTTGTGCTTACAGC 1311  
QY 661 ATGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACACATTTGGATTTAGCAGATC 720  
Db 1312 ATGAGCTGAGCCAAACCATCCGAAATACAGATGGAAGGACACATTTGGATTTAGCAGATC 1371  
QY 721 CATCTGCCAAAGCAGTCTTACTGTTGAATATAAGAAAGATCAACTTTAGAAAGTGCCA 780  
Db 1372 CATCTGCCAAAGCAGTCTTACTGTTGAATATAAGAAAGATCAACTTTAGAAAGTGCCA 1431  
QY 781 GGAGTGGCAATGAAGAAAAATGATGGCTCTACTCACACCATTAATGTCAACTGCCAGC 840  
Db 1432 GGAGTGGCAATGAAGAAAAATGATGGCTCTACTCACACCATTAATGTCAACTGCCAGC 1491  
QY 841 CAAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGA 900  
Db 1492 CAAGTATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGA 1551  
QY 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTTAAAGATAAAGCTGATCTGG 960  
Db 1552 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTTAAAGATAAAGGTGATCTGG 1611  
QY 961 TACCATTACACAATGCTGTCTTATGGTCATTATGAAGTAAGTCAACTGAACTTTTGGTCAAGC 1020  
Db 1612 TACCATTACACAATGCTGTCTTATGGTCATTATGAAGTAAGTCAACTGAACTTTTGGTCAAGC 1671  
QY 1021 ATGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTATGAGGAGCAGCTT 1080  
Db 1672 ATGGTGCCTGTGTAATGCAATGGACTTGTGGCAATTCACCTCTTATGAGGAGCAGCTT 1731  
QY 1081 CTAAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGC 1140  
Db 1732 CTAAAGAACAGGTTGAAGTATGTTCTCTCTTAAAGTTATGGTGCAGACCCCAACACTGC 1791  
QY 1141 TCAATTGTCAATAAAAGTGTATAGACTTGGCTCCCAACACAGATTAAAGAAAGAT 1200  
Db 1792 TCAATTGTCAATAAAAGTGTATAGACTTGGCTCCCAACACAGATTAAAGAAAGAT 1851  
QY 1201 TAGCATATGAATTTAAAGGCGCTCTGCTGCAAGCTGCAGAGAGCTGATGTTACTGC 1260  
Db 1852 TAGCATATGAATTTAAAGGCGCTCTGCTGCAAGCTGCAGAGAGCTGATGTTACTGC 1911

QY 1261 GAATCAAAAAACATCTCTCTCGAAATGGTGAATTTTCAAGCATCCTCAAAACACATGAAA 1320  
Db 1912 GAATCAAAAAACATCTCTCTCGAAATGGTGAATTTTCAAGCATCCTCAAAACACATGAAA 1971  
QY 1321 CAGCATTCGATTTGCTGCTGCATCTCCATATCCAAAAGAAAGCAAAATATGTGAACGTG 1380  
Db 1972 CAGCATTCGATTTGCTGCTGCATCTCCATATCCAAAAGAAAGCAAAATATGTGAACGTG 2031  
QY 1381 TGCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAGAATTTCTTGACTCCTCTGACAG 1440  
Db 2032 TGCTAAGAAAAGGAGCAACATCAATGAAAAGACTAAGAATTTCTTGACTCCTCTGACAG 2091  
QY 1441 TGGCATCTGAAAAGCTCATATATGATGTTTGAAGTAGTGGTGAACATCAAGCAAGG 1500  
Db 2092 TGGCATCTGAAAAGCTCATATATGATGTTTGAAGTAGTGGTGAACATCAAGCAAGG 2151  
QY 1501 TTAATGCTCTCGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 1560  
Db 2152 TTAATGCTCTCGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTTGGTGCATC 2211  
QY 1561 TACAACCTTCGCCCTACTCTGAGCTATGGTGTGATCCTTAACATTTATATCCCTTCAGG 1620  
Db 2212 TACAACCTTCGCCCTACTCTGAGCTATGGTGTGATCCTTAACATTTATATCCCTTCAGG 2271  
QY 1621 GCTTTACTGCTTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGSTATCT 1680  
Db 2272 GCTTTACTGCTTTTACAGATGGAAATGAAATGTACAGCAACTCTCCAGAGGSTATCT 2331  
QY 1681 CATTAGSTTAATTCAGAGGACAGACACAATTTGCTGGAGCTGCAAGGCTGGAGATGTCG 1740  
Db 2332 CATTAGSTTAATTCAGAGGACAGACACAATTTGCTGGAGCTGCAAGGCTGGAGATGTCG 2391  
QY 1741 AAACGTAAAAAAACTGTGTACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 1800  
Db 2392 AAACGTAAAAAAACTGTGTACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCGTC 2451  
QY 1801 AGTCTACACCACTTCATTTTGCAGCTGGTATAACAGAGTGTCCGTGGTGAATATCTGC 1860  
Db 2452 AGTCTACACCACTTCATTTTGCAGCTGGTATAACAGAGTGTCCGTGGTGAATATCTGC 2511  
QY 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTGTACCTTTGACACAATG 1920  
Db 2512 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCTTGTACCTTTGACACAATG 2571  
QY 1921 CATGTTCTTATGGACATTTATGAAGTTGCAGAACCTTCTTGTAAACATGGAGCAGTAGTTA 1980  
Db 2572 CATGTTCTTATGGACATTTATGAAGTTGCAGAACCTTCTTGTAAACATGGAGCAGTAGTTA 2631  
QY 1981 ATGTAGCTGATTTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAATATG 2040  
Db 2632 ATGTAGCTGATTTTATGGAAATTTACACCTTTTACATGAAGCAGCAGCAAAAGAAATATG 2691  
QY 2041 AAAATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTTACCACAAAAAACAAGGATGAA 2100  
Db 2692 AAAATTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTTACCACAAAAAACAAGGATGAA 2751  
QY 2101 ATACTCCTTTTGGATCTCTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAG 2160  
Db 2752 ATACTCCTTTTGGATCTCTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAG 2811  
QY 2161 ATGACGCTTTTCTAGATCTGCCAAGAGGCTTGTTTTAGCCAGAGTGAAGAAGTTGTCTT 2220  
Db 2812 ATGACGCTTTTCTAGATCTGCCAAGAGGCTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2871  
QY 2221 CTCCTGATAATGTAAATTTGCCCGCATACCCCAAGGCAGACATTTCAACACCTTTACATTTAG 2280  
Db 2872 CTCCTGATAATGTAAATTTGCCCGCATACCCCAAGGCAGACATTTCAACACCTTTACATTTAG 2931  
QY 2281 CAGCTGGTTATTAATTAATTTAGAAAGTTGCAGAGTATTTGTTTACAACACGGAGCTGATGTA 2340  
Db 2932 CAGCTGGTTATTAATTAATTTAGAAAGTTGCAGAGTATTTGTTTACAACACGGAGCTGATGTA 2991



```
QY 2341 ATCCCAAGCAAAAGGAGGACTTATTCCTTTACATATAGCAGCATCTTACGGGCATGTAG 2400
Db 2992 ATCCCAAGCAAAAGGAGGACTTATTCCTTTACATATAGCAGCATCTTACGGGCATGTAG 3051
QY 2401 ATGTAGCAGCTTACTTAATAAAGTATAATGCGATGTGTCATGTCACCGGACAAATGGGCTT 2460
Db 3052 ATGTAGCAGCTTACTTAATAAAGTATAATGCGATGTGTCATGTCACCGGACAAATGGGCTT 3111
QY 2461 TCACACCTTTGACGAAGCAGCCAAAGGAGGAGCAACACAGCTTTTGTGCTTTGCTGCTAG 2520
Db 3112 TCACACCTTTGACGAAGCAGCCAAAGGAGGAGCAACACAGCTTTTGTGCTTTGCTGCTAG 3171
QY 2521 CCATGGAGCTGACCCGACTCTTTAAATAATCAGGAAGCAAAACACCTTTAGATTAGTTT 2580
Db 3172 CCATGGAGCTGACCCGACTCTTTAAATAATCAGGAAGCAAAACACCTTTAGATTAGTTT 3231
QY 2581 CAGCGATGATGTCAGGGCTCTCTCAGCAGCAGCCATGCCCCATCTGCTCTGCCCTCTT 2640
Db 3232 CAGCAGATGATGTCAGGGCTCTCTCAGCAGCAGCCATGCCCCATCTGCTCTGCCCTCTT 3291
QY 2641 GTTACAAGCTCAAGTGCTCAATGGTGTGAGAAAGCCAGGAGCACTGCAGATGCTCTCT 2700
Db 3292 GTTACAAGCTCAAGTGCTCAATGGTGTGAGAAAGCCAGGAGCACTGCAGATGCTCTCT 3351
QY 2701 CTTTCAGGTCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGCTTTGACAACTTATCTG 2760
Db 3352 CTTTCAGGTCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGCTTTGACAACTTATCTG 3411
QY 2761 GGAGTTTTTCAGAACTCTCTCAGTAGTAGTTTCAAGTGAAGGAGGCTTCCAGTT 2820
Db 3412 GGAGTTTTTCAGAACTCTCTCAGTAGTAGTTTCAAGTGAAGGAGGCTTCCAGTT 3471
QY 2821 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATTAAGTCACTCAATTCGTAAGGAATCTTG 2880
Db 3472 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATTAAGTCACTCAATTCGTAAGGAATCTTG 3531
QY 2881 GACTTCAGCACCTTAATGGATATATTTAGAGAGAAACAGATCACTTTGGATGTATTAGTTG 2940
Db 3532 GACTTCAGCACCTTAATGGATATATTTAGAGAGAAACAGATCACTTTGGATGTATTAGTTG 3591
QY 2941 AGATGGGCACAAAGAGCTGAAGGAGATGGAATCAATCTGATGACATAGGCACAAAC 3000
Db 3592 AGATGGGCACAAAGAGCTGAAGGAGATGGAATCAATCTGATGACATAGGCACAAAC 3061
QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGGCAACAAGGCTTTAACCCATATTAACTT 3060
Db 3652 TAATTAAGGAGTCGAGAGACTTATCTCCGGCAACAAGGCTTTAACCCATATTAACTT 3711
QY 3061 TGAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCCTGATGATAAAGAGTTTC 3120
Db 3712 TGAACACCTCTGCTAGTGGAACTTCTTATAGATCTGCTCCTGATGATAAAGAGTTTC 3771
QY 3121 AGTCTGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGAGATGAGGTTCATGCGAG 3180
Db 3772 AGTCTGTGGAGGAGAGATGCAAGTACAGTTTCGAGAGCAGAGATGAGGTTCATGCGAG 3831
QY 3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAAGAGTTTCTTAACAAGAACTAT 3240
Db 3832 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAAGAGTTTCTTAACAAGAACTAT 3891
QY 3241 GGGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCACTGCGCAATG 3300
Db 3892 GGGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCACTGCGCAATG 3951
QY 3301 AACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATATTCACAAGGCTTTGATG 3360
Db 3952 AACGAATGCTATTTTCATGGGTCTCTTTTGTGAATGCAATATTCACAAGGCTTTGATG 4011
QY 3361 AAAGCATCGGTACATAGTGGTGTATTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 3420
Db 4012 AAAGCATCGGTACATAGTGGTGTATTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 4071
QY 3421 CCAAAAGCAATCAATATGTATATGGAAATGGAGGAGGTACTGGGTGTCAGATTTCACAAAG 3480
|||||
Db 4072 CCAAAAGCAATCAATATGTATATGGAAATGGAGGAGGTACTGGGTGTCAGATTTCACAAAG 4131
QY 3481 ACAGATCTTGTTCACATTTGCCACAGCAGCTGCTCTTTTCCCGGTAACCTTGGAAAGT 3540
Db 4132 ACAGATCTTGTTCACATTTGCCACAGCAGCTGCTCTTTTCCCGGTAACCTTGGAAAGT 4191
QY 3541 CTTTCCCTGCAGTTTCAGTGGCAATGAAATGGCACATTTCTCTCCAGGTCATCACTCAGTCA 3600
Db 4192 CTTTCCCTGCAGTTTCAGTGGCAATGAAATGGCACATTTCTCTCCAGGTCATCACTCAGTCA 4251
QY 3601 CTGGTAGGCCAGCTGTAATGCGCTAGCATTAGCTGAATATGTTATTACAGAGGAGAAC 3660
Db 4252 CTGGTAGGCCAGCTGTAATGCGCTAGCATTAGCTGAATATGTTATTACAGAGGAGAAC 4311
QY 3661 AGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTATGAGGCTGAGGTATGGTCGATG 3720
Db 4312 AGCTTATCTCTGAGTATTTAAATTTACTTACCAGATTATGAGGCTGAGGTATGGTCGATG 4371
QY 3721 GATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTG 3780
Db 4372 GATAAATAGTTATTTTAAAGAACTAATTTCCACTGAACCTTAAATCATCAAGCAGCAGTG 4431
QY 3781 GCCTCTACGTTTCTACTCTTGTGTAAGAAAAA 3813
Db 4432 GCCTCTACGTTTCTACTCTTGTGTAAGAAAAA 4464

RESULT 6
AAF63952
ID AAF63952 standard; DNA; 5002 BP.
XX
AC AAF63952;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-LONG coding sequence SEQ ID NO: 132.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
XX
PR 29-JUN-1999; 99US-0141582.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR WPI; 2001-102896/11.
XX
P-PSDB; AAB66294.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX
PS Claim 5; Page 185-190; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
```



QY 2101 ATACTCCTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTCAGATCTCTTAGGGAG 2160  
DB 2762 ATACTCCTTTGGATCTTGTGTTAAAGATGGAGATACAGATATTCAGATCTCTTAGGGAG 2821  
QY 2161 ATCAGCTTTGCTAGATGCTGCCAAGAAGGTTGTTTAGCCAGAGTGAAGAAGTGTCTTT 2220  
DB 2822 ATCAGCTTTGCTAGATGCTGCCAAGAAGGTTGTTTAGCCAGAGTGAAGAAGTGTCTTT 2881  
QY 2221 CTCTGATATGTAATGTTGCGCGGATACCCAAAGCGACATTCACACACCTTTACATTTAG 2280  
DB 2882 CTCTGATATGTAATGTTGCGCGGATACCCAAAGCGACATTCACACCTTTACATTTAG 2941  
QY 2281 CAGCTGGTTATATAATTTAGAGTTGTCAGAGTATTTGTTTACAAACGCGAGCTGATGTGA 2340  
DB 2942 CAGCTGGTTATATAATTTAGAGTTGTCAGAGTATTTGTTTACAAACGCGAGCTGATGTGA 3001  
QY 2341 ATCCCAAGCAAGGAGGACTTATTCCTTTACATATATGATGTCATATGCGGCGATGTAG 2400  
DB 3002 ATCCCAAGCAAGGAGGACTTATTCCTTTACATATATGATGTCATATGCGGCGATGTAG 3061  
QY 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTCATATGCGGCGACAAATGGGCTT 2460  
DB 3062 ATGTAGCAGCTCTACTAATAAAGTATAATGATGTCATATGCGGCGACAAATGGGCTT 3121  
QY 2461 TCACACCTTTGCGAGAGAGCCCAAGAGGAGACACAGCTTTGTGCTTTGCTAG 2520  
DB 3122 TCACACCTTTGCGAGAGAGCCCAAGAGGAGACACAGCTTTGTGCTTTGCTAG 3181  
QY 2521 CCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGACAAACACCTTTAGATTTAGTTT 2580  
DB 3182 CCATGGAGCTGACCCGACTCTTTAAAAATCAGGAAGACAAACACCTTTAGATTTAGTTT 3241  
QY 2581 CAGCGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640  
DB 3242 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3301  
QY 2641 GTTACAGGCTCAAGTGCTCAATGGTGTGAGAGCCAGAGGCTGAGATGCTCTCT 2700  
DB 3302 GTTACAGGCTCAAGTGCTCAATGGTGTGAGAGCCAGAGGCTGAGATGCTCTCT 3361  
QY 2701 CTTGAGGTCATCTAGCCCATCAAGCCTTTCTGACGCGAGCAGCTGTGACAACTTATCTG 2760  
DB 3362 CTTGAGGTCATCTAGCCCATCAAGCCTTTCTGACGCGAGCAGCTGTGACAACTTATCTG 3421  
QY 2761 GGAGTTTTCAGAACTGCTTCAGTAGTTAGTTCAAGTGGAAAGAGGCTGCTTCCAGTT 2820  
DB 3422 GGAGTTTTCAGAACTGCTTCAGTAGTTAGTTCAAGTGGAAAGAGGCTGCTTCCAGTT 3481  
QY 2821 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTAGCATACTCAATTCGTAAGGAATCTTG 2880  
DB 3482 TGGAGAAAAGGAGGTTCCAGGAGTAGATTTAGCATACTCAATTCGTAAGGAATCTTG 3541  
QY 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGATTAAGTTG 2940  
DB 3542 GACTTGAGCACCTAATGGATATATTTGAGAGAGACAGATCACTTTGGATGATTAAGTTG 3601  
QY 2941 AGATGGGACAAAGAGCTGAAGGAGATTGGAATCAATGCTTATGACATAGSCAACAC 3000  
DB 3602 AGATGGGACAAAGAGCTGAAGGAGATTGGAATCAATGCTTATGACATAGSCAACAC 3661  
QY 3001 TAATTAAGGAGTCGAGAGACTTATCTCCGACACAAAGGCTTAAACCATATTTAACTT 3060  
DB 3662 TAATTAAGGAGTCGAGAGACTTATCTCCGACACAAAGGCTTAAACCATATTTAACTT 3721  
QY 3061 TGAACACCTCTGGTAGTGAACAATTTATAGATCTGCTCTGATGATAAAGAGTTTC 3120  
DB 3722 TGAACACCTCTGGTAGTGAACAATTTATAGATCTGCTCTGATGATAAAGAGTTTC 3781  
QY 3121 AGTCTGTGAGAGAGATGCAAGATACAGTTTCGAGAGACAGATGAGGTCATGCGAG 3180  
DB 3782 AGTCTGTGAGAGAGATGCAAGATACAGTTTCGAGAGACAGATGAGGTCATGCGAG 3841  
QY 3181 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAAGAGTTTGTCAACAAGAACTAT 3240

DB 3842 GTGGAATCTTCAACAGATACAATATTTCTCAAGATTCAAGAGTTTGTFAACAAGAACTAT 3901  
QY 3241 GSGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCATGCCAATG 3300  
DB 3902 GSGAAAGATACACTCACCGGAGAAAAGAGTTTCTGAAGAAAACCAACCATGCCAATG 3961  
QY 3301 AACGAATGCTTATTTCAATGGGTCTCTTTTGTGAATGCAATATTCACAAAAGGCTTTCATG 3360  
DB 3962 AACGAATGCTTATTTCAATGGGTCTCTTTTGTGAATGCAATATTCACAAAAGGCTTTCATG 4021  
QY 3361 AAAGGATCGGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 3420  
DB 4022 AAAGGATCGGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTGTAAGAACTCTT 4081  
QY 3421 CCAAAAGCAATCAATATGATATGGAATGGAGAGGAGTACTGGGTGTCAGTTTCACAAAG 3480  
DB 4082 CCAAAAGCAATCAATATGATATGGAATGGAGAGGAGTACTGGGTGTCAGTTTCACAAAG 4141  
QY 3481 ACAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTCCCGGGTAACCTTGGGAAAGT 3540  
DB 4142 ACAGATCTTGTACATTTGCCACAGCAGCTGCTCTTTTCCCGGGTAACCTTGGGAAAGT 4201  
QY 3541 CTTTCTCGAGTTTCAAGTGAATGAAATGGACATCTCTCTCCAGGTCATCTCACTCACTCA 3600  
DB 4202 CTTTCTCGAGTTTCAAGTGAATGAAATGGACATCTCTCTCCAGGTCATCTCACTCACTCA 4261  
QY 3601 CTGCTAGGCCCCAGTGAATGGCTAGCAATAGCTAGTGAATGTAATTTTACAGAGGAGAAC 3660  
DB 4262 CTGCTAGGCCCCAGTGAATGGCTAGCAATAGCTAGTGAATGTAATTTTACAGAGGAGAAC 4321  
QY 3661 AGGCTTATCTCGAGTATTTAAATTTACTTACCAGATTTAGGCGCTGAAGGTATGTCGATG 3720  
DB 4322 AGGCTTATCTCGAGTATTTAAATTTACTTACCAGATTTAGGCGCTGAAGGTATGTCGATG 4381  
QY 3721 GATAAATAGTATTTTAAAGAACTAATTCACCTGAACCTTAAATATCATCAAGCAGCAGTG 3780  
DB 4382 GATAAATAGTATTTTAAAGAACTAATTCACCTGAACCTTAAATATCATCAAGCAGCAGTG 4441  
QY 3781 GCCTCTACCTTTTACTCTTGTGAAAAAAA 3813  
DB 4442 GCCTCTACCTTTTACTCTTGTGAAAAAAA 4474

RESULT 7  
AAF63926  
ID AAF63926 standard; DNA; 4127 Bp.  
XX AAF63926;  
XX  
XX  
DT 05-APR-2001 (first entry)  
XX Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 100.  
DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
KW inflammatory disorder; ds.  
XX Homo sapiens.  
XX OS  
XX WO200100849-A1.  
XX  
XX PD 04-JAN-2001.  
XX  
PF 28-JUN-2000; 2000WO-US17827.  
XX  
PR 29-JUN-1999; 99US-0141582.  
XX  
PA (ICOS-) ICOS CORP.  
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
PI WPI; 2001-102896/11.  
XX P-PSDB; AAB66288.  
DR

XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
 PT inflammatory and autoimmune disorders -  
 XX  
 XX Example 1; Page 158-162; 242pp; English.

XX The present invention provides the protein and coding sequence for the  
 CC human tankyrase2 protein. This is found in two different versions,  
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
 CC polyADP-ribosylation activity and is involved in the modification of  
 CC TRF1, which is a telomere-specific binding protein. The regulation of  
 CC telomere length, in which TRF1 has a role, is linked to ageing and  
 CC cancer. The sequences are useful in the treatment of cancers and  
 CC inflammatory disorders.

XX Sequence 4127 BP; 1245 A; 813 C; 934 G; 1135 T; 0 other:

Query Match 94.2%; Score 3595.8; DB 22; Length 4127;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3597; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 215 GGCCAGGATCATGTGGGTGCGCGTGGCGGGGGAGCGGCGTGGCGAGCGCGC 274

Db 1 GGCCAGGATCATGTGGGTGCGCGTGGCGGGGGAGCGGCGTGGCGAGCGCGC 60

QY 275 GGCCAGGCGGTGGAGCGCGCGCCGAGAGCTGTTGAGGCGTGGCGACGCGGCGT 334

Db 61 GGCCAGGCGGTGGAGCGCGCGCCGAGAGCTGTTGAGGCGTGGCGACGCGGAGCT 120

QY 335 GGAACGAGTCAAGAGCGTGGTACGCCGTGAGAGGTGAACGCCGACGCGGCGGAG 394

Db 121 GGAACGAGTCAAGAGCGTGGTACGCCGTGAGAGGTGAACGCCGACGCGGCGGAG 180

QY 395 GAAATCCACCGCGTGCATCTGGCGCAGGTTTGGCGGAAAGACGCTAGTTGAATATT 454

Db 181 GAAATCCACCGCGTGCATCTGGCGCAGGTTTGGCGGAAAGACGCTAGTTGAATATT 240

QY 455 GCTTCAGAAATGTCAAATGTCAGACGCTGATGATGGGGCGCTTATTCCTCTTCATAA 514

Db 241 GCTTCAGAAATGTCAAATGTCAGACGCTGATGATGGGGCGCTTATTCCTCTTCATAA 300

QY 515 TGCATGCTCTTTTGGTGCATGTGAAGTAGTCAATCTCCTTTTGGACATGGTCAGACCC 574

Db 301 TGCATGCTCTTTTGGTGCATGTGAAGTAGTCAATCTCCTTTTGGACATGGTCAGACCC 360

QY 575 CAATGCTCGAGATATTGGAATTATCTCCCTCCATGAAGCTGCAATTAAGGAAAGAT 634

Db 361 CAATGCTCGAGATATTGGAATTATCTCCCTCCATGAAGCTGCAATTAAGGAAAGAT 420

QY 635 TGATGTTTGCATGTGCTTTACAGCATGGAGCTGAGCAACCATCCGAAATACAGATGG 694

Db 421 TGATGTTTGCATGTGCTTTACAGCATGGAGCTGAGCAACCATCCGAAATACAGATGG 480

QY 695 AAGGACAGCATTTGATTTAGCAGATCCATCTGCCAAGCAGTGGCTTACTGGTGAATATA 754

Db 481 AAGGACAGCATTTGATTTAGCAGATCCATCTGCCAAGCAGTGGCTTACTGGTGAATATA 540

QY 755 GAAAGATGAACCTTTAGAAAGTCCAGAGTGGCAATGAAGAAAAATGATGGCTCTACT 814

Db 541 GAAAGATGAACCTTTAGAAAGTCCAGAGTGGCAATGAAGAAAAATGATGGCTCTACT 600

QY 815 CACACCATTTAAATGTCACTGCCAGCAGTGTGGCAGAAAGTCAACTCCATTCATATT 874

Db 601 CACACCATTTAAATGTCACTGCCAGCAGTGTGGCAGAAAGTCAACTCCATTCATATT 660

QY 875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTGTACTGCAACATGGAGCTGATGT 934

Db 661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTGTACTGCAACATGGAGCTGATGT 720

QY 935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACATGCTCTTCTTATGGTCATTA 994

Db 721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACATGCTCTTCTTATGGTCATTA 780

QY 995 TGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGCACTTGTGGCA 1054

Db 781 TGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGCACTTGTGGCA 840

QY 1055 ATTCACTCTCTTTTCATGAGGAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTTCTCTT 1114

Db 841 ATTCACTCTCTTTTCATGAGGAGCTTCTTAAGAACAGGGTTGAAGTATGTTCTTCTCTT 900

QY 1115 AGTTTATGGTGCAGAGCCCAACACTGCTCAATTTGTCAATAAAGTGTATAGACTTGGC 1174

Db 901 AGTTTATGGTGCAGAGCCCAACACTGCTCAATTTGTCAATAAAGTGTATAGACTTGGC 960

QY 1175 TCCACACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234

Db 961 TCCACACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020

QY 1235 AGCTGCACGAGAAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTCGAATGTTGAA 1294

Db 1021 AGCTGCACGAGAAGCTGATGTTTACTCGAATCAAAAAACATCTCTCTCGAATGTTGAA 1080

QY 1295 TTTCAAGCATCTCAACACATCAACAGCATTTGCATTTGCTGTGCTGCTATCCCATATCC 1354

Db 1081 TTTCAAGCATCTCAACACATCAACAGCATTTGCATTTGCTGTGCTGCTATCCCATATCC 1140

QY 1355 CAAAAAGAACCAATATGTGAACCTGTTGCTAAGAAAGAGCAACATCAATGAAAAAGAC 1414

Db 1141 CAAAAAGAACCAATATGTGAACCTGTTGCTAAGAAAGAGCAACATCAATGAAAAAGAC 1200

QY 1415 TAAAGAAATTTTGAACCTCTGACACGCTGCAATCTGAGAAAGCTCAATATGATGTTTGA 1474

Db 1201 TAAAGAAATTTTGAACCTCTGACACGCTGCAATCTGAGAAAGCTCAATATGATGTTTGA 1260

QY 1475 AGTAGTGTCAACATGAACCAAGGTTAATGCTCTGGAATACTTTGGTCAGACTTCTCT 1534

Db 1261 AGTAGTGTCAACATGAACCAAGGTTAATGCTCTGGAATACTTTGGTCAGACTTCTCT 1320

QY 1535 ACACAGAGCTGCAATATTTGCTGCTATCTACAAAGCTGCGGCTACTCTCTGAGCTATGGGTG 1594

Db 1321 ACACAGAGCTGCAATATTTGCTGCTATCTACAAAGCTGCGGCTACTCTCTGAGCTATGGGTG 1380

QY 1595 TGATCTTAACTTATATCTCTTCCAGGCTTTACTGCTTTTACAGATGGGAAATGAAATGT 1654

Db 1381 TGATCTTAACTTATATCTCTTCCAGGCTTTACTGCTTTTACAGATGGGAAATGAAATGT 1440

QY 1655 ACAGCAACTCTCTCAAGAGGTTATCTCATTAGGTAAATTCAGAGCGACAGACAATTTGCT 1714

Db 1441 ACAGCAACTCTCTCAAGAGGTTATCTCATTAGGTAAATTCAGAGCGACAGACAATTTGCT 1500

QY 1715 GGAAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTACAGAGTGT 1774

Db 1501 GGAAGCTGCAAGAGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTACAGAGTGT 1560

QY 1775 CAATGTCAGACATTTCAAGGGGCTGCTGCTACACCACTTCAATTTTGCAGCTGGGTATAA 1834

Db 1561 CAATGTCAGACATTTCAAGGGGCTGCTGCTACACCACTTCAATTTTGCAGCTGGGTATAA 1620

QY 1835 CAGAGTGTCCGTGGTGAATATCTGCTACAGATGGAGCTGATGCTGATGCTAAAGATAA 1894

Db 1621 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGCTGATGCTAAAGATAA 1680

QY 1895 AGGAGGCTTGTACCTTTGCACAATGCTGCTTTATGGACATTTAAGTTGCGAGAACT 1954

Db 1681 AGGAGGCTTGTACCTTTGCACAATGCTGCTTTATGGACATTTAAGTTGCGAGAACT 1740

QY 1955 TCTTTTAAACATGGAGCAGTAGTAAATGATGCTGATTTATGGAATTTACACCTTTTACA 2014

Db 1741 TCTTTTAAACATGGAGCAGTAGTAAATGATGCTGATTTATGGAATTTTACACCTTTTACA 1800

QY 2015 TGAAGCAGCAGCAAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGA 2074

Db 1801 TGAAGCAGCAGCAAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGA 1860

QY 2075 CCCTACCAAAAAAACAGGGATGGAATACCTTTGGATCTTGTAAAGATGGAGATAC 2134  
Db 1861 CCCACAAAAAAACAGGGATGGAATACCTTTGGATCTTGTAAAGATGGAGATAC 1920  
QY 2135 AGATATTCAAGATCTGCTTAGGGAGATGACGCTTTGCTAGATGCTGCCAAGAAGGTTG 2194  
Db 1921 AGATATTCAAGATCTGCTTAGGGAGATGACGCTTTGCTAGATGCTGCCAAGAAGGTTG 1980  
QY 2195 TTTAGCCAGAGTGAAGAAGTTGCTTCTCCCTGATAATGTTAAATTCGCCGCGATACCCRAAG 2254  
Db 1981 TTTAGCCAGAGTGAAGAAGTTGCTTCTCCCTGATAATGTTAAATTCGCCGCGATACCCRAAG 2040  
QY 2255 CAGACATTCAACACCTTTACACTTTAGCAGCTGGTTATAATAATTTAGAAAGTTGAGAGTA 2314  
Db 2041 CAGACATTCAACACCTTTACACTTTAGCAGCTGGTTATAATAATTTAGAAAGTTGAGAGTA 2100  
QY 2315 TTTCTTACACACAGGAGCTGATGTGAATGCCCCAGACAAAGGAGGACTTATTCCTTTACA 2374  
Db 2101 TTTCTTACACACAGGAGCTGATGTGAATGCCCCAGACAAAGGAGGAGCTTATTCCTTTACA 2160  
QY 2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGCATG 2434  
Db 2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTATAAAGTATAATGCATG 2220  
QY 2435 TGTCAAATGCCACGACAAATGGGCTTTACACCTTTGCACGAAGCAGCCCAAAAGGAGC 2494  
Db 2221 TGTCAAATGCCACGACAAATGGGCTTTACACCTTTGCACGAAGCAGCCCAAAAGGAGC 2280  
QY 2495 AACACAGCTTTGCTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554  
Db 2281 AACACAGCTTTGCTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340  
QY 2555 AGGACAAACACCTTTAGATTAGTTTTCAGCGGATGTGTGAGCGCTCTTCTGCAGCAGC 2614  
Db 2341 AGGACAAACACCTTTAGATTAGTTTTCAGCGAGATGTGTGAGCGCTCTTCTGCAGCAGC 2400  
QY 2615 CATGCCCCCATCTGCTCTGCCCTCTGTTACAGCCCTCAAGTGTCAATGGTGTGAAG 2674  
Db 2401 CATGCCCCCATCTGCTCTGCCCTCTGTTACAGCCCTCAAGTGTCAATGGTGTGAAG 2460  
QY 2675 CCCAGGAGCCACTGCAGATGCTCTCTCCAGGTGCCATCTAGCCCATCAAGCCCTTCTGC 2734  
Db 2461 CCCAGGAGCCACTGCAGATGCTCTCTCCAGGTGCCATCTAGCCCATCAAGCCCTTCTGC 2520  
QY 2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCAGTAGTTTC 2794  
Db 2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTCTTTCAGAACTGTCTTCAGTAGTTTC 2580  
QY 2795 AAGTGGAACAGAGGCTGCTTCCAGTTTGGAGAAAGAGGAGGTTCCAGGAGTAGATTTAG 2854  
Db 2581 AAGTGGAACAGAGGCTGCTTCCAGTTTGGAGAAAGAGGAGGTTCCAGGAGTAGATTTAG 2640  
QY 2855 CATAACTCAATTCGTAGGAATCTTGGACTTGGACCTTAATGGATATATTTCAGAGAGA 2914  
Db 2641 CATAACTCAATTCGTAGGAATCTTGGACTTGGACCTTAATGGATATATTTCAGAGAGA 2700  
QY 2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGCAACAAGGAGCTGAAGGAGATTGGAAT 2974  
Db 2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGCAACAAGGAGCTGAAGGAGATTGGAAT 2760  
QY 2975 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGACA 3034  
Db 2761 CAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGCTTATCTCCGACA 2820  
QY 3035 ACAAGTCTTAACCCATATTAACTTTGAACACCTCTGCTAGTGGAACTTCTTTATAGA 3094  
Db 2821 ACAAGTCTTAACCCATATTAACTTTGAACACCTCTGCTAGTGGAACTTCTTTATAGA 2880  
QY 3095 TCTGCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAGAGATGCCAAGTACAGTTG 3154  
Db 2881 TCTGCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAGAGATGCCAAGTACAGTTG 2940  
QY 3155 AGAGCACAGATGGAGGTTCATGCAGTGGATCTTCAACAGATCAATATCTCAAGAT 3214

Db 2941 AGACACAGAGATGGAGTCTATGAGGTGGATCTTCAACAGATACAATATCTCAAGAT 3000  
QY 3215 TCAGAAGTTTGTAAACAAGAACTATGGAAAGATACACTCACCGGAGAAAGATTTTC 3274  
Db 3001 TCAGAAGTTTGTAAACAAGAACTATGGAAAGATACACTCACCGGAGAAAGATTTTC 3060  
QY 3275 TGAAGAAACACACACCAATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTGTGAA 3334  
Db 3061 TGAAGAAACACACACCAATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTGTGAA 3120  
QY 3335 TGCAATTATCCACAAAGGCTTTGATGAAAGCATGCGTACATAGTGTATGTTGGAGC 3394  
Db 3121 TGCAATTATCCACAAAGGCTTTGATGAAAGCATGCGTACATAGTGTATGTTGGAGC 3180  
QY 3395 TGGCATTATTTTCTGCTGAAAACCTCTTCCAAAGCAATCAATATGATGCAATTTGGAGG 3454  
Db 3181 TGGCATTATTTTCTGCTGAAAACCTCTTCCAAAGCAATCAATATGATGCAATTTGGAGG 3240  
QY 3455 AGGTACTGGGTGTCAGTTCCAGTTCAAAAGCAGATCTTGTACATTTGCCACAGCAGCTGCT 3514  
Db 3241 AGGTACTGGGTGTCAGTTCCAGTTCAAAAGCAGATCTTGTACATTTGCCACAGCAGCTGCT 3300  
QY 3515 CTTTTCGCCGGTAACCTTTGGGAAAGTCTTCTCCTGCAGTTCAAGTCAATGAAAATGGCACA 3574  
Db 3301 CTTTTCGCCGGTAACCTTTGGGAAAGTCTTCTCCTGCAGTTCAAGTCAATGAAAATGGCACA 3360  
QY 3575 TTCTCCTCCAGGTCATCAGTCAGTCACTGGTAGGCCAGTGTAAATGSCCTAGCATTTAGC 3634  
Db 3361 TTCTCCTCCAGGTCATCAGTCAGTCACTGGTAGGCCAGTGTAAATGSCCTAGCATTTAGC 3420  
QY 3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGCAGTTCAAGTCAATGAAAATGGCACA 3694  
Db 3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGCAGTTCAAGTCAATGAAAATGGCACA 3480  
QY 3695 TATGAGCCCTCAAGGATGTTGATGGATAAAATAGTTATTTTAAAGAACTAAATTTCCACTG 3754  
Db 3481 TATGAGCCCTCAAGGATGTTGATGGATAAAATAGTTATTTTAAAGAACTAAATTTCCACTG 3540  
QY 3755 AACCTAAATCATCAAGACAGCAGTGGCCCTAGCTTTTACTCCTTTGCTGAAAAAAA 3813  
Db 3541 AACCTAAATCATCAAGACAGCAGTGGCCCTAGCTTTTACTCCTTTGCTGAAAAAAA 3599

RESULT 8  
AAD02578  
ID AAD02578 standard; DNA; 4512 Bp.  
XX AAD02578;  
XX 02-MAY-2001 (first entry)  
XX Human tankyrase homolog protein (THP) DNA.  
DE Human; tankyrase homolog protein (THP) DNA.  
XX  
KW Human; tankyrase homolog protein; THP; gene therapy; cancer;  
KW tumour; basal cell carcinoma; therapy; genetic mapping;  
KW cytostatic; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
CDS 23..3523  
FT /\*tag= a  
FT /product= "Human tankyrase homolog protein (THP)"  
FT /note= "This region is specifically claimed in  
FT claim 4 as SEQ ID NO:4"  
XX  
PN WO200104326-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 03-JUL-2000; 2000WO-BP06609.  
XX



QY	1830	TATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGTCATGCTAAA	1889
DB	1628	TATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGTCATGCTAAA	1687
QY	1890	GATAAAGGAGGCCCTTGTTACCTTTTGCACAATGCATGTTCTTTATGGACATATATGAAGTTGCA	1949
DB	1688	GATAAAGRRGCCCTTGTTACCTTTTGCACAATGCATGTTCTTTATGGACATATATGAAGTTGCA	1747
QY	1950	GAACCTCTCTGTTAAACATGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTTACACCT	2009
DB	1748	GAACCTCTCTGTTAAACATGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTTACACCT	1807
QY	2010	TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT	2069
DB	1808	TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT	1867
QY	2070	GCAGACCCTACCAAAAAACACAGGATGAAATACTCCTTTTGGATCTTGTTTAAGATGGA	2129
DB	1868	GCAGACCCTACCAAAAAACACAGGATGAAATACTCCTTTTGGATCTTGTTTAAGATGGA	1927
QY	2130	GATACAGATATTCGAAGATCTGCTTAGGGGAGATGCAGCTTTCGTAGATGCTGCCAAGAAG	2189
DB	1928	GATACAGATATTCGAAGATCTGCTTAGGGGAGATGCAGCTTTCGTAGATGCTGCCAAGAAG	1987
QY	2190	GTTTGTGTAGCCAGTGAAGAAGTTGTCTTCTCTGATAATGTAAATTTGCCGATACC	2249
DB	1988	AGTTTGTGTAGCCAGTGAAGAAGTTGTCTTCTCTGATAATGTAAATTTGCCGATACC	2047
QY	2250	CAAGCAGACATTCACACCTTTTACATTTTAGCAGCTGGTTATATTAATTTAGAAGTTGCA	2309
DB	2048	CAAGCAGACATTCACACCTTTTACATTTTAGCAGCTGGTTATATTAATTTAGAAGTTGCA	2107
QY	2310	GAGTATTTGTTTACAACACGGAGCTGATGTAATGCCAAGACAAAGGAGGACTTTATTCCT	2369
DB	2108	GAGTATTTGTTTACAACACGGAGCTGATGTAATGCCAAGACAAAGGAGGACTTTATTCCT	2167
QY	2370	TTACATTAATGCAGATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGATATAAT	2429
DB	2168	TTACATTAATGCAGATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGATATAAT	2227
QY	2430	GCATGTCTCAATGCCAGCACAAATGGGCTTTTCACACCTTTGCAGGAAGACGCCAAAAG	2489
DB	2228	GCATGTCTCAATGCCAGCACAAATGGGCTTTTCACACCTTTGCAGGAAGACGCCAAAAG	2287
QY	2490	GGAGAACACAGCTTTGTGCTTTGTCTTAGGCCATGGAGCTGACCCGACTCCTTAAAAAT	2549
DB	2288	GGAGAACACAGCTTTGTGCTTTGTCTTAGGCCATGGAGCTGACCCGACTCCTTAAAAAT	2347
QY	2550	CAGGAGGACAAACACCTTTAGTATTTTACGGGATGATGTACGCGCTCTTCTGACA	2609
DB	2348	CAGGAGGACAAACACCTTTAGTATTTTACGGGATGATGTACGCGCTCTTCTGACA	2407
QY	2610	GCAGCCATGCCCCATCTGCTCTGCCCTTGTTTACAGCCCTCAAGTGTCTCAATGGTGTG	2669
DB	2408	GCAGCCATGCCCCATCTGCTCTGCCCTTGTTTACAGCCCTCAAGTGTCTCAATGGTGTG	2467
QY	2670	AGAAGCCACGAGGCCACTGCATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGCCTT	2729
DB	2468	AGAAGCCACGAGGCCACTGCATGCTCTCTTTCAGGTCCATCTAGCCCATCAAGCCTT	2527
QY	2730	TTCTGAGCCAGATCTTGACAACTTATCTGGAGTTTTTCAGAACTGTCTTTCAGTAGTT	2789
DB	2528	TTCTGAGCCAGATCTTGACAACTTATCTGGAGTTTTTCAGAACTGTCTTTCAGTAGTT	2587
QY	2790	AGTTCAAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGAT	2849
DB	2588	AGTTCAAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGAT	2647
QY	2850	TTTATGATTAACCTCAATTCGTAAGGAATCTTGACTTGAGCACCTTAATGGATATATTGAG	2909
DB	2648	TTTATGATTAACCTCAATTCGTAAGGAATCTTGACTTGAGCACCTTAATGGATATATTGAG	2707
QY	2910	AGAGAACAGATCACTTTTGGATGTATTAGTTGAGATGGGGCAAGAGGAGCTTGAAGGAGATT	2969

Db	2708	AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGCAAGGAGCTCAAGGAGATT	2767
QY	2970	GGAAATCAATGCTTATGGACATAGCCACAACTAATTAAGAGGTCGAGAGACATTATCTCC	3029
Db	2768	GGAAATCAATGCTTATGGACATAGGCACAACTAATTAAGAGGTCGAGAGACATTATCTCC	2827
QY	3030	GGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGCTAGTGGAAACAATCTTT	3089
Db	2828	GGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGCTAGTGGAAACAATCTTT	2887
QY	3090	ATAGATCTGTCCTCTGATGATTAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA	3149
Db	2888	ATAGATCTGTCCTCTGATGATTAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA	2947
QY	3150	GTTCGAGACACAGATGGAGGTCATGCAGGTGGAACTTTCAACAGATACAATATTTCTC	3209
Db	2948	GTTCGAGACACAGATGGAGGTCATGCAGGTGGAACTTTCAACAGATACAATATTTCTC	3007
QY	3210	AAGATTCAAGGTTTGTTAACAAGAACTATGGGAAGATACACTCACCCGGAGAAAAGAA	3269
Db	3008	AAGATTCAAGGTTTGTTAACAAGAACTATGGGAAGATACACTCACCCGGAGAAAAGAA	3067
QY	3270	GTTCCTGAAGAAAACACAAACCATGCAATGAACGAATGCTATTTCATGGTCTCCTTTT	3329
Db	3068	GTTCCTGAAGAAAACACAAACCATGCAATGAACGAATGCTATTTCATGGTCTCCTTTT	3127
QY	3330	GTGAATGCAATTATCCACAAAGGCTTTGATGAAGAGCATGCGTACATAGTGGTATGTATT	3389
Db	3128	GTGAATGCAATTATCCACAAAGGCTTTGATGAAGAGCATGCGTACATAGTGGTATGTATT	3187
QY	3390	GGAGCTGGCAATTATTTTCTGANAACCTCTTCCAAAGCAATCAATATGTATATGGAATT	3449
Db	3188	GGAGCTGGCAATTATTTTCTGANAACCTCTTCCAAAGCAATCAATATGTATATGGAATT	3247
QY	3450	GGAGGAGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG	3509
Db	3248	GGAGGAGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG	3307
QY	3510	CTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCCTTCTCGAGTTCAGTCAATGAAATG	3569
Db	3308	CTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCCTTCTCGAGTTCAGTCAATGAAATG	3367
QY	3570	GCACATTCTCCTCCAGGTCATCACCTCAGTCAGTGTAGGCCCAGTGAATGGCCTAGCA	3629
Db	3368	GCACATTCTCCTCCAGGTCATCACCTCAGTCAGTGTAGGCCCAGTGAATGGCCTAGCA	3427
QY	3630	TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAAATTACTTAC	3689
Db	3428	TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAAATTACTTAC	3487
QY	3690	CAGATTATGAGCCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAAGAACTAATTC	3749
Db	3488	CAGATTATGAGCCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAAGAACTAATTC	3547
QY	3750	CACCTCAACCTTAAATCATCAAGCAGCAGTGGCCTCTACGTTTTTACTCCTTTGCTCAAAA	3809
Db	3548	CACCTCAACCTTAAATCATCAAGCAGCAGTGGCCTCTACGTTTTTACTCCTTTGCTCAAAA	3607
QY	3810	AAAA 3813	
Db	3608	AAAA 3611	

RESULT 9	
AAF63837	
ID	AAF63837 standard; DNA; 3508 BP.
XX	
AC	AAF63837;
XX	
XX	
DT	05-APR-2001 (first entry)
XX	
DE	Human tankyrase2 related coding sequence SEQ ID NO: 1.





QY 1775 CAACTGCAGAGACATTGAAGGCGCTCAGCTTACACCACCTTCATTTTCAGCTGGGTATAA 1834  
Db 1561 CAACTGCAGAGACATTGAAGGCGCTCAGCTTACACCACCTTCATTTTCAGCTGGGTATAA 1620  
QY 1835 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894  
Db 1621 CAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680  
QY 1895 AGGAGGCTTGTACCTTTGCACATGCTCTCTTATGGACATATGAAGTTGCAGAACT 1954  
Db 1681 AGGAGGCTTGTACCTTTGCACATGCTCTCTTATGGACATATGAAGTTGCAGAACT 1740  
QY 1955 TCTTGTAAACATGGAGCAGTAGTAAATAGTGTGATTTATGAAATTTACACCTTTACA 2014  
Db 1741 TCTTGTAAACATGGAGCAGTAGTAAATAGTGTGATTTATGAAATTTACACCTTTACA 1800  
QY 2015 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGA 2074  
Db 1801 TGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGAGA 1860  
QY 2075 CCCTACCAAAAAACAGGAGTGAATATCTCTCTTGGATCTTGTAAAGATGGAGATAC 2134  
Db 1861 CCCTACCAAAAAACAGGAGTGAATATCTCTTGGATCTTGTAAAGATGGAGATAC 1920  
QY 2135 AGATATTCAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGTTG 2194  
Db 1921 AGATATTCAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCCAAGAGGTTG 1980  
QY 2195 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTCTGATATGTAATTTAGAACTTTCAGAGTA 2254  
Db 1981 TTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATTTAGAACTTTCAGAGTA 2040  
QY 2255 CAGACATTCAACACCTTTACATTTAGCAGCTGTTTATAATTTAGAACTTTCAGAGTA 2314  
Db 2041 CAGACATTCAACACCTTTACATTTAGCAGCTGTTTATAATTTAGAACTTTCAGAGTA 2100  
QY 2315 TTTGTTTACACAGGAGCTGATGTGAATGCCAAGCAAGAGGAGGACTTATCTCTTTACA 2374  
Db 2101 TTTGTTTACACAGGAGCTGATGTGAATGCCAAGCAAGAGGAGGACTTATCTCTTTACA 2160  
QY 2375 TAATGAGCATCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAAGTATAATCATG 2434  
Db 2161 TAATGAGCATCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAAGTATAATCATG 2220  
QY 2435 TGTCAATGCCAGGACAATGGCTTTACACCTTTGCAGAGCAGGCCCAAAAGGAGC 2494  
Db 2221 TGTCAATGCCAGGACAATGGCTTTACACCTTTGCAGAGCAGGCCCAAAAGGAGC 2280  
QY 2495 AACACAGCTTTGTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554  
Db 2281 AACACAGCTTTGTGCTTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340  
QY 2555 AGGACAAACACCTTTAGTATTACGGGATGTAGTGTGAGCTCTTCTGACAGCAGC 2614  
Db 2341 AGGACAAACACCTTTAGTATTACGGGATGTAGTGTGAGCTCTTCTGACAGCAGC 2400  
QY 2615 CATGCCCATCTGCTCTGCCCTCTGTTTACAGGCTCAAGTGTCAATGGTGTGAGAG 2674  
Db 2401 CATGCCCATCTGCTCTGCCCTCTGTTTACAGGCTCAAGTGTCAATGGTGTGAGAG 2460  
QY 2675 CCCAGGAGCCTGACAGATGCTCTCTTTCAGGTCATCTAGCCCATCAAGCTTCTGTC 2734  
Db 2461 CCCAGGAGCCTGACAGATGCTCTCTTTCAGGTCATCTAGCCCATCAAGCTTCTGTC 2520  
QY 2735 AGCCAGAGCTTTGACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794  
Db 2521 AGCCAGAGCTTTGACAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580  
QY 2795 AAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAGAGGTTTCAGGAGTAGATTTTAG 2854  
Db 2581 AAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAGAGGTTTCAGGAGTAGATTTTAG 2640  
QY 2855 CATAACTCAATTCGTAAAGAACTCTTGGACTTGAACACCTAATGGATATATTGAGAGAGA 2914

Db 2641 CATAACTCAATTCGTAAAGAACTTTGGACTTGAACACCTAATGGATATATTGAGAGAGA 2700  
QY 2915 ACAGATCACCTTTGATGTATTAGTTGAGATGGGCGCAAAAGGAGCTGAAGGAGATTGAAT 2974  
Db 2701 ACAGATCACCTTTGATGTATTAGTTGAGATGGGCGCAAAAGGAGCTGAAGGAGATTGAAT 2760  
QY 2975 CAATGCTTATGGACATAGGCACAACTTAATTAAGAGAGTCGAGAGACTTATCTCCGGACA 3034  
Db 2761 CAATGCTTATGGACATAGGCACAACTTAATTAAGAGAGTCGAGAGACTTATCTCCGGACA 2820  
QY 3035 ACAAGGCTTTAAACCAATATTAACTTTGAACACCTCTGGTAGTGGAACTTCTTATAGA 3094  
Db 2821 ACAAGGCTTTAAACCAATATTAACTTTGAACACCTCTGGTAGTGGAACTTCTTATAGA 2880  
QY 3095 TCTGTCTCTGATGATAAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGC 3154  
Db 2881 TCTGTCTCTGATGATAAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAAAGTACAGTTGC 2940  
QY 3155 AGAGCACAGAGATGGAGTTCATGCAGTGGAACTTCTCAACAGATACAAATTTCTCAAGAT 3214  
Db 2941 AGAGCACAGAGATGGAGTTCATGCAGTGGAACTTCTCAACAGATACAAATTTCTCAAGAT 3000  
QY 3215 TCAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGGAGAAAGAGTTTC 3274  
Db 3001 TCAGAAGGTTTGTAAACAAGAACTATGGGAAAGATACACTCACCGGAGAAAGAGTTTC 3060  
QY 3275 TGAAGAAACCAACACATGCCAATGAACGAATGCTATTTTCATGGGTCCTCTTTGTGAA 3334  
Db 3061 TGAAGAAACCAACACATGCCAATGAACGAATGCTATTTTCATGGGTCCTCTTTGTGAA 3120  
QY 3335 TGCAATTTATCCCAAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTGGAGC 3394  
Db 3121 TGCAATTTATCCCAAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTGGAGC 3180  
QY 3395 TGGCATTTATTTGCTGAAACCTCTTCCAAAGCAATCAATATGTATATGGAATTTGGAG 3454  
Db 3181 TGGCATTTATTTGCTGAAACCTCTTCCAAAGCAATCAATATGTATGGAATTTGGAG 3240  
QY 3455 AGGTACTGGTGTCCAGTTCACAAAGACAGATCTTGTACATTTGCCACAGCAGCTGCT 3514  
Db 3241 AGGTACTGGTGTCCAGTTCACAAAGACAGATCTTGTACATTTGCCACAGCAGCTGCT 3300  
QY 3515 CTTTTCGCCGGTAACTTGGGAAAGTCTTTCCTGCAAGTTCAGTGCATGAAATGACACA 3574  
Db 3301 CTTTTCGCCGGTAACTTGGGAAAGTCTTTCCTGCAAGTTCAGTGCATGAAATGACACA 3360  
QY 3575 TTCTCCTCAGGTCACTCACTAGTGGTGGCCCAAGTGTAAATGGCCTAGCATTAGC 3634  
Db 3361 TTCTCCTCAGGTCACTCACTAGTGGTGGCCCAAGTGTAAATGGCCTAGCATTAGC 3420  
QY 3635 TGAATATGTTATTTACAGAGGAGACAGGCTTATCCTGAGTATTTAACTTACCAGAT 3694  
Db 3421 TGAATATGTTATTTACAGAGGAGACAGGCTTATCCTGAGTATTTAACTTACCAGAT 3480  
QY 3695 TATGAGGCTGAAGGTATGTCGATGGA 3722  
Db 3481 TATGAGGCTGAAGGTATGTCGATGGA 3508

## RESULT 10

AAA91487

ID AAA91487 standard; DNA; 3797 bp.

XX AAA91487;

AC AAA91487;

XX 06-AUG-2001 (first entry)

DT Tankyrase homologue isotype 1 coding sequence.

DE Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;

XX cell cycle protein; cell cycle associated disorder; cancer; gene mapping;

KW chromosome mapping; gene therapy; vaccine; ds.

KW

```

XX OS Unidentified.
XX FH Key Location/Qualifiers
XX CDS 404..3706
XX FT /*tag= a
XX FT /partial
XX FT /product= "TaHoI"
XX FT /note= "tanyrase homologue isotype 1; No start
XX FT codon given"
XX PN WO200130987-A2.
XX PD 03-MAY-2001.
XX XX 25-OCT-2000; 2000WO-US41528.
XX PF 25-OCT-1999; 99US-0427154.
XX PR (RIGE-) RIGEL PHARM INC.
XX PA
XX XX Luo Y, Chan E, Xu X, Huang B;
XX PI WPI; 2001-300503/31.
XX DR P-PSDB; AAY97748.
XX XX
XX PT Novel recombinant cell cycle polypeptide, tankyrase H useful for
XX PT inducing or preventing cell proliferation in cells, and for diagnosing,
XX PT treating or preventing cell cycle associated disorders such as cancer
XX PT .
XX PS Claim 13; Fig 1; 63pp; English.
XX CC This sequence encodes the Tankyrase homologue isotype 1 (TaHo-1) protein
XX CC of the invention. The invention also relates to the TaHo-2 protein.
XX CC The TaHo proteins are useful for inducing or preventing cell
XX CC proliferation in cells, and in the study or treatment of conditions
XX CC mediated by the cell cycle proteins, such as to diagnose, treat or
XX CC prevent cell cycle associated disorders, preferably cancer. The TaHo
XX CC coding sequences are useful as hybridisation probes, in chromosome and
XX CC gene mapping and in the generation of anti-sense DNA and RNA. The coding
XX CC sequences are also useful for the preparation of TaHo, for generating
XX CC either transgenic animals or knock out animals which, in turn, are useful
XX CC in a development and screening of therapeutically useful agents, in gene
XX CC therapy, as vaccine, and for construction of hybridisation probes for
XX CC mapping the gene which encodes TaHo and for the genetic analysis of
XX CC individuals with genetic disorders. The TaHo proteins, and their coding
XX CC sequences are useful in screening assays.
XX SQ Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 other;

Query Match 88.9%; Score 3393.4; DB 22; Length 3797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 422 AGGTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAAATGGTGCAATGTCCTCAAGC 481
DB 403 AGGTTTGGCGGAAAGAGCTAGTTGAATATTTGCTTCAGAAATGGTGCAATGTCCTCAAGC 462
QY 482 AGGTGATGATGGGGGCTTATTCCTTCATTAATGCAATGCTCTTTTGGTCATGCTGAAGT 541
DB 463 AGGTGATGATGGGGGCTTATTCCTTCATTAATGCAATGCTCTTTTGGTCATGCTGAAGT 522
QY 542 AGTCAATCTCCTTTTGGACATGGTGCAGACCCCAATGCTGCAGATAATTTGAATTTATAC 601
DB 523 AGTCAATCTCCTTTTGGACATGGTGCAGACCCCAATGCTGCAGATAATTTGAATTTATAC 582
QY 602 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTTGATGTTTGCATTTGCTGTTACAGCA 661
DB 583 TCCTCTCCATGAAGCTGCAATTTAAAGGAAAGATTTGATGTTTGCATTTGCTGTTACAGCA 642
QY 662 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAGGACAGCATTTGGATTTAGCAGATCC 721
DB 1723 AACTGTAAAAAACTGTGTACTGTTTACAGATGTTCACTGCAGAGACATTTGAAGGCGCTCA 1801
DB 1723 AACTGTAAAAAACTGTGTACTGTTTACAGATGTTCACTGCAGAGACATTTGAAGGCGCTCA 1782

```

```

DB 643 TGGAGCTGAGCCCAACCATCCGAAATACAGATGGAAGCAGCAGATTGGATTTAGCAGATCC 702
QY 722 ATCTGCCAAGCAGTCTTACTTGGTCAATATAGAAAGATGAACCTTTAGAAAGTCCAG 781
DB 703 ATCTGCCAAGCAGTCTTACTTGGTCAATATAGAAAGATGAACCTTTAGAAAGTCCAG 762
QY 782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
DB 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
QY 842 AAGTATGCGCAAGTCAACTCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 901
DB 823 AAGTATGCGCAAGTCAACTCCATTACATTTGGCAGCAGGATATACAGAGTAAAGAT 882
QY 902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGTAAAGATAAAGTGTATGCTGT 961
DB 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGTAAAGATAAAGTGTATGCTGT 942
QY 962 ACCATTACACAATGCCCTGTTCTTATGGTCAATATGAAGTAACCTGAACTTTTGGTCAAGCA 1021
DB 943 ACCATTACACAATGCCCTGTTCTTATGGTCAATATGAAGTAACCTGAACTTTTGGTCAAGCA 1002
QY 1022 TGGTGCCTGTGTAATGCAATGGAGCTTGGGCAATTCACCTCTTTCATGAGCAGCTTC 1081
DB 1003 TGGTGCCTGTGTAATGCAATGGAGCTTGGGCAATTCACCTCTTTCATGAGCAGCTTC 1062
QY 1082 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 1141
DB 1063 TAAGAACAGGGTTGAAGTATGTTCTTCTTAAAGTTATGGTGCAGACCCCAACACTGCT 1122
QY 1142 CAATTGTACAAATAAAGTGCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1201
DB 1123 CAATTGTACAAATAAAGTGCTATAGACTTGGCTCCACACACAGTTAAAGAAAGATT 1182
QY 1202 AGCATATGAATTTAAAGGCCACACTGTTGCTGCAAGCTGCACAGAGAGCTGATCTTACTCG 1261
DB 1183 AGCATATGAATTTAAAGGCCACACTGTTGCTGCAAGCTGCACAGAGAGCTGATCTTACTCG 1242
QY 1262 AATCAAAAAACATCTCTCTCGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAAC 1321
DB 1243 AATCAAAAAACATCTCTCTCGGAAATGGTGAATTTCAAGCATCCTCAACACATGAAAC 1302
QY 1322 AGCATTCATTTGCTGCTGCTGCATCTCCATATCCAAAGAAAGCAATATGTGAACCTGTT 1381
DB 1303 AGCATTCATTTGCTGCTGCTGCATCTCCATATCCAAAGAAAGCAATATGTGAACCTGTT 1362
QY 1382 GCTAAGAAAGAGGCAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGCACGT 1441
DB 1363 GCTAAGAAAGAGGCAACATCAATGAAAGACTTAAAGAAATTTCTTGACTCCTCTGCACGT 1422
QY 1442 GGCACTCTGAGAAAGCTCATATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1501
DB 1423 GGCACTCTGAGAAAGCTCATATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGT 1482
QY 1502 TAATGCTCTGATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGGTGCATCT 1561
DB 1483 TAATGCTCTGATAATCTTGGTCAGACTTCTCTACAGAGCTGCATATTTGGTGCATCT 1542
QY 1562 ACAACCTCGCGGCTTACTCCTCTGAGCTATGGGTGTGATCTTAAACATTTATCCCTTCAAGG 1621
DB 1543 ACAACCTCGCGGCTTACTCCTCTGAGCTATGGGTGTGATCTTAAACATTTATCCCTTCAAGG 1602
QY 1622 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCAAGAGGATATCTC 1681
DB 1603 CTTTACTGCTTTACAGATGGGAAATGAAATGTACAGCAACTCCTCAAGAGGATATCTC 1662
QY 1682 ATTAGGTAAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGAGATGTCGA 1741
DB 1663 ATTAGGTAAATTCAGAGGACAGACACAATTTGCTGGAAGCTGCAAGGCTGAGATGTCGA 1722
QY 1742 AACTGTAAAAAACTGTGTACTGTTTACAGATGTTCACTGCAGAGACATTTGAAGGCGCTCA 1801
DB 1723 AACTGTAAAAAACTGTGTACTGTTTACAGATGTTCACTGCAGAGACATTTGAAGGCGCTCA 1782

```







```
QY 2941 AGATGGGGCACAGAGCTGAAGGAGATTGGAAATCAATGCTTATGACATAGGCACACAC 3000
DB 3010 AGATGGGGCACAGAGCTGAAGGAGATTGGAAATCAATGCTTATGACATAGGCACACAC 3069
QY 3001 TAATTAAGAGAGTCGAGAGACTTATCTCGGACACAAAGGCTTAAACCCATATTTAACTT 3060
DB 3070 TAATTAAGAGTCGAGAGACTTATCTCGGACACAAAGGCTTAAACCCATATTTAACTT 3129
QY 3061 TGAACACCTCTGTAGTGAACAATTTCTTATAGATCTGCTCTGATGATAAAGAGTTTC 3120
DB 3130 TGAACACCTCTGTAGTGAACAATTTCTTATAGATCTGCTCTGATGATAAAGAGTTTC 3189
QY 3121 AGTCTGTGGAGGAAGAGATGCAAGTACAGTTCGAGAGCACAGAGATGAGGTTCATGCAG 3180
DB 3190 AGTCTGTGGAGGAAGAGATGCAAGTACAGTTCGAGAGCACAGAGATGAGGTTCATGCAG 3249
QY 3181 GTGGAATCTTCAACAGATACAATTTCTCAAGATTTCAGAAAGTTTCTTAAC-----A 3231
DB 3250 GTGGAATCTTCAACAGATACAATTTCTCAAGATTTCAGAAAGTTTCTTAACAGAGCCCAAGA 3309
QY 3232 AGAACTATGGGAAAGATACACTCACGGGAGAAAAGAGTTTCTGAAGAAAACCAACAC 3291
DB 3310 TTCCGGACGAGGAAAGATACACTCACGGGAGAAAAGAGTTTCTGAAGAAAACCAACAC 3369
QY 3292 ATGCAATGAACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCACAAAG 3351
DB 3370 ATGCAATGAACGAATGCTATTTTCATGGGTCTCCTTTTGTGAATGCAATTTATCCACAAAG 3429
QY 3352 GCTTTGATGAAGAGATGCGTACATAGTGGTATGTTTGGAGCTGCAATTTATTTGGTGTG 3411
DB 3430 GCTTTGATGAAGAGATGCGTACATAGTGGTATGTTTGGAGCTGCAATTTATTTGGTGTG 3489
QY 3412 AAAACTCTTCAAAAGCAATCAATATGATATGGAATGGAGAGGTACTGGGTGTCAG 3471
DB 3490 AAAACTCTTCAAAAGCAATCAATATGATATGGAATGGAGAGGTACTGGGTGTCAG 3549
QY 3472 TTCACAAAGACAGATCTGTTACATTTGCCACAGGAGCTGCTCTTTTGGCGGGTAACCT 3531
DB 3550 TTCACAAAGACAGATCTGTTACATTTGCCACAGGAGCTGCTCTTTTGGCGGGTAACCT 3609
QY 3532 TGGGAAAGTCTTCTCGAGTTCAGTGCATGCAATGAAATGGCACATTTCTCTCCAGGTCAATC 3591
DB 3610 TGGGAAAGTCTTCTCGAGTTCAGTGCATGCAATGAAATGGCACATTTCTCTCCAGGTCAATC 3669
QY 3592 ACTCAGTCACTGGTAGCCAGTGAATGCCTAGCATTAGCTGAATATGTTATTTACA 3651
DB 3670 ACTCAGTCACTGGTAGCCAGTGAATGCCTAGCATTAGCTGAATATGTTATTTACA 3729
QY 3652 GAGGAGAACAGGCTTATCCTGAGTATTTAATTTACTTACCAGATTATGAGGCTGAAGGTA 3711
DB 3730 GAGGAGAACAGGCTTATCCTGAGTATTTAATTTACTTACCAGATTATGAGGCTGAAGGTA 3789
QY 3712 TGGTCGATGGATAAATAGTTATTTTATAGAACTAATTTCCACTGAACCTAAATCATCAA 3771
DB 3790 TGGTCGATGGATAAATAGTTATTTTATAGAACTAATTTCCACTGAACCTAAATCATCAA 3849
QY 3772 GCAGCAGTGGCCTCTAGCTTTTACTCTCTTCTGTAAGAAAAA 3813
DB 3850 GCAGCAGTGGCCTCTAGCTTTTACTCTCTTCTGTAAGAAAAA 3891

RESULT 12
AA25366
ID AA25366 standard; cDNA; 3400 BP.
XX
AC
XX
XX
DT 19-JUL-1999 (first entry)
XX
DE Human Grb7 effector 2.2412 cDNA.
XX
KW Grb7 effector; 2.2412 protein; human; signal transduction;
tumour marker; breast cancer; prostate cancer; prognosis;
```

```
KW diagnosis; ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 1..3225
FT CDS /*tag= a
XX
PN W09915647-A1.
XX
PD 01-APR-1999.
XX
PF 23-SEP-1998; 98WO-AU00795.
XX
PR 23-SEP-1997; 97AU-0009388.
XX
PA (GARV-) GARVAN INST MEDICAL RES.
XX
PI Daly RJ, Sutherland RL;
XX
DR WPI: 1999-254707/21.
XX
P-PSDB; AAY05734.
XX
PT New candidate effector for the Grb7 family of signaling proteins,
PS and specific antibody, useful for detection and treatment of cancer
XX
PS Claim 4; Fig 1; 24pp; English.
XX
CC This is the nucleotide sequence of an isolated polynucleotide
CC molecule encoding a candidate effector protein, termed 2.2412 (see
CC AAY05734), for the Grb7 family of signalling proteins. A partial
CC clone was isolated from a human liver cDNA library using a yeast
CC two-hybrid system assay with Grb14 as bait. This clone was used
CC to screen a human placental cDNA library that provided the
CC C-terminal end of the 2.2412 sequence. A 5' sequence for the
CC clone, including the initiation codon, was not identified.
CC Analysis of the sequence revealed significant homology to a large
CC number of proteins containing ankyrin-like repeats. The gene
CC was localised to between chromosome 10q23.2 and proximal 10q23.32.
CC Deletions in the 10q22-25 region have been detected in human breast,
CC prostate, renal, small cell lung and endometrial carcinomas.
CC glioblastoma multiforme, melanoma and meningiomas. Detection of
CC the protein encoded by the 2.2412 cDNA in a sample should provide a
CC useful tumour marker and/or prognostic indicator for certain human
CC cancers, in particular breast cancer and prostate cancer.
CC Antagonism of the interaction between Grb7 family members and the
CC encoded protein should provide a novel treatment strategy for human
CC diseases exhibiting aberrant receptor tyrosine kinase signalling,
CC such as cancer. Oligonucleotide probes can be used in methods of
CC detecting the presence of 2.2412 mRNA in a sample.
XX
SQ Sequence 3400 BP; 1059 A; 665 C; 760 G; 916 T; 0 other;
```

```
Query Match 86.4%; Score 3297; DB 20; Length 3400;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 501 ATTCTCTTCAATCAATGCTCTTTTGGTGCATGCTGGAAGTAGTCAATCTCTTTTGGCA 560
DB 1 ATTCTCTTCAATCAATGCTCTTTTGGTGCATGCTGGAAGTAGTCAATCTCTTTTGGCA 60

QY 561 CATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATCTCTCTCCATGAAGCTGCA 620
DB 61 CATGGTGCAGACCCCAATGCTCGAGATAATTTGGAATTTATCTCTCTCCATGAAGCTGCA 120

QY 621 ATTAAAGGAAGATTGATGTTTGCATTTGCTCTTACAGCATGGAGTCAACCATC 680
DB 121 ATTAAAGGAAGATTGATGTTTGCATTTGCTCTTACAGCATGGAGTCAACCATC 180

QY 681 CGAAATACAGATGGAAGGACAGCATTGGATTATAGCAGATCCATCTGCCAAAGCAGTCTT 740
DB 181 CGAAATACAGATGGAAGGACAGCATTGGATTATAGCAGATCCATCTGCCAAAGCAGTCTT 740
```



QY 741 ACTGGTGAATATAGAAGATGAACCTTTAGAAAGTCCAGGAGTGGCAATGAAGAAAA 800  
Db 241 ACTGGTGAATATAGAAGATGAACCTTTAGAAAGTCCAGGAGTGGCAATGAAGAAAA 300  
QY 801 ATGATGGCTCTACTCACACCATTAATGTCACCTGCGACGCAAGTGAAGGAGAGAGTCA 860  
Db 301 ATGATGGCTCTACTCACACCATTAATGTCACCTGCGACGCAAGTGAAGGAGAGAGTCA 360  
QY 861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGATTGTACAGCTGTACTGCA 920  
Db 361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAGATTGTACAGCTGTACTGCA 420  
QY 921 CATGGAGCTGATGTCATGCTTAAGATAAAGGTGATCTGGTACCATTAACAAATGCTGT 980  
Db 421 CATGGAGCTGATGTCATGCTTAAGATAAAGGTGATCTGGTACCATTAACAAATGCTGT 480  
QY 981 TCTATAGTCAATTAAGTAACTGAACCTTTGGTCAAGCATGGTGGCTGTGTAATGCA 1040  
Db 481 TCTATAGTCAATTAAGTAACTGAACCTTTGGTCAAGCATGGTGGCTGTGTAATGCA 540  
QY 1041 ATGGACTTTGGCAATTCACCTCTTCATGAGGCAAGTCTTAAGAACAGGTTGAAGTA 1100  
Db 541 ATGGACTTTGGCAATTCACCTCTTCATGAGGCAAGTCTTAAGAACAGGTTGAAGTA 600  
QY 1101 TGTCTCTCTCTTAAGTATGGTGAGACCCCAACACTGCTCAATTTGCACAAATAAAGT 1160  
Db 601 TGTCTCTCTCTTAAGTATGGTGAGACCCCAACACTGCTCAATTTGAAGATAAAGT 660  
QY 1161 GCTATAGACTTGGCTCCCAACACACAGTTTAAAGAAAGATTAGCATATCAATTTAAGGC 1220  
Db 661 GCTATAGACTTGGCTCCCAACACACAGTTTAAAGAAAGATTAGCATATCAATTTAAGGC 720  
QY 1221 CACTCGTGTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAACATCTCTCT 1280  
Db 721 CACTCGTGTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAACATCTCTCT 780  
QY 1281 CTGGAATAGGTCAATTTCAAGCATCTCAACACATGAACAGCATGTCATTTGCTGTCT 1340  
Db 781 CTGGAATAGGTCAATTTCAAGCATCTCAACACATGAACAGCATGTCATTTGCTGTCT 840  
QY 1341 GCATCTCCATATCCCAAGAAAGCAATATATGTAACCTGTTGCTAAGAAAGAGCAAC 1400  
Db 841 GCATCTCCATATCCCAAGAAAGCAATATGTAACCTGTTGCTAAGAAAGAGCAAC 900  
QY 1401 ATCAATGAAGACATAAGAAATTTTGACTCTCTGACGTGGCATCTGAGAAAGCTCAT 1460  
Db 901 ATCAATGAAGACATAAGAAATTTTGACTCTCTGACGTGGCATCTGAGAAAGCTCAT 960  
QY 1461 AATGATGTTGTTGAAGTAGTGTGAACATGAAGCAAGGTTAATGCTCTGGATATCTT 1520  
Db 961 AATGATGTTGTTGAAGTAGTGTGAACATGAAGCAAGGTTAATGCTCTGGATATCTT 1020  
QY 1521 GGTACAGCTTCTTACAGAGCTGCATATTTGGTCAATCTACAAACCTGCGCCCTACTC 1580  
Db 1021 GGTACAGCTTCTTACAGAGCTGCATATTTGGTCAATCTACAAACCTGCGCCCTACTC 1080  
QY 1581 CTGAGCTATGGTGTGATCTTAACATTTATCCCTTACAGGCTTTTACGCTTTACAGATG 1640  
Db 1081 CTGAGCTATGGTGTGATCTTAACATTTATCCCTTACAGGCTTTTACGCTTTACAGATG 1140  
QY 1641 GGAATGAAATGTACAGCAACTCTCTCAAGAGGATATCTCATTTAGTAAATTCAGAGCA 1700  
Db 1141 GGAATGAAATGTACAGCAACTCTCTCAAGAGGATATCTCATTTAGTAAATTCAGAGCA 1200  
QY 1701 GACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGT 1760  
Db 1201 GACAGCAATTTGCTGGAAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGT 1260  
QY 1761 ACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCACTTCATTTT 1820  
Db 1261 ACTGTTACAGTGTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCACTTCATTTT 1320  
QY 1821 GCAGCTGGGTATAACAGAGTGTCTGCTGGTGAATATCTGCTACAGCATGGAGCTGATGTG 1880

Db 1321 GCAGCTGGGTATACAGAGTGTCCGTGGTGAATATCTGTCAGCATGGAGCTGATGTG 1380  
QY 1881 CATGCTAAAGATAAAGAGGCGCTTGACCTTTGCACAATGCATGTTCTTTATGGACATTAT 1940  
Db 1381 CATGCTAAAGATAAAGAGGCGCTTGACCTTTGCACAATGCATGTTCTTTACGGACATTAT 1440  
QY 1941 GAAGTTGCAGAACTTCTTCTTTAAACATGAGCAGTAGTATTAATGTAGCTGATTTATGAAA 2000  
Db 1441 GAAGTTGCAGAACTTCTTCTTTAAACATGAGCAGTAGTATTAATGTAGCTGATTTATGAAA 1500  
QY 2001 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAATTTGCAAACTTCTGCTC 2060  
Db 1501 TTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAATTTGCAAACTTCTGCTC 1560  
QY 2061 CAGCATGTTGTCAGACCTTACCACCAAAAAACAGGATGGAATACTCCTTTGGATCTTGT 2120  
Db 1561 CAGCATGTTGTCAGACCTTACCACCAAAAAACAGGATGGAATACTCCTTTGGATCTTGT 1620  
QY 2121 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2180  
Db 1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680  
QY 2181 GCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAAAGTCTCTCTCTGATATGTAATTTGC 2240  
Db 1681 GCCAAGAGGGTGTGTTTAGCCAGAGTGAAGAAAGTCTCTCTCTGATATGTAATTTGC 1740  
QY 2241 CCGGATACCCAGGAGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATATAATTTA 2300  
Db 1741 CCGGATACCCAGGAGCAGACATTCACACCTTTACATTTAGCAGCTGGTTATATAATTTA 1800  
QY 2301 GAAGTTGCAGAGTATTTGTACACACGAGCTGATGTAATGCCCAAGCAAGGAGGA 2360  
Db 1801 GAAGTTGCAGAGTATTTGTACACACGAGCTGATGTAATGCCCAAGCAAGGAGGA 1860  
QY 2361 CTTATTCCTTTACATTAATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420  
Db 1861 CTTATTCCTTTACATTAATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920  
QY 2421 AAGTATATGTCATGTCAATGCCAGGACAATGGGCTTTTACACCTTTGACAGCAAGA 2480  
Db 1921 AAGTATATGTCATGTCTCAATGCCAGGACAATGGGCTTTTACACCTTTGACAGCAAGA 1980  
QY 2481 GCCCAAGAGGAGCAACACAGCTTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACT 2540  
Db 1981 GCCCAAGAGGAGCAACACAGCTTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACT 2040  
QY 2541 CTTTAAATTCAGGAAGGACAAACACCTTTTAGATTTAGTTTACGGGATGATCTCAGCGCT 2600  
Db 2041 CTTTAAATTCAGGAAGGACAAACACCTTTTAGATTTAGTTTACAGCAGATGATCTCAGCGCT 2100  
QY 2601 CTTCTGACAGAGCCATGCCCATCTGCTGCTGCTCTTTGTATACAGCTCAAGTGCTC 2660  
Db 2101 CTTCTGACAGAGCCATGCCCATCTGCTGCTGCTCTTTGTATACAGCTCAAGTGCTC 2160  
QY 2661 AATGTTGTGAGAAGCCCAAGGAGCCACTGCAGATGCTCTCTCTCAGGTCATCTAGCCCA 2720  
Db 2161 AATGTTGTGAGAAGCCCAAGGAGCCACTGCAGATGCTCTCTCTCAGGTCATCTAGCCCA 2220  
QY 2721 TCAAGCCCTTCTGACGCGCAGCTTTGACAACTTATCTGGAGTGTTCAGAACTGTCT 2780  
Db 2221 TCAAGCCCTTCTGACGCGCAGCTTTGACAACTTATCTGGAGTGTTCAGAACTGTCT 2280  
QY 2781 TCAGTAGTTAGTTCAAGTGAACAGAGGCTCTCCACTTTGGAGAAAAGAGGTTCCA 2840  
Db 2281 TCAGTAGTTAGTTCAAGTGAACAGAGGCTCTCCACTTTGGAGAAAAGAGGTTCCA 2340  
QY 2841 GGAGTAGATTTTACGATAACTCAATTCGTAAGGAATCTTGGACTTTGACACCTAATGGAT 2900  
Db 2341 GGAGTAGATTTTACGATAACTCAATTCGTAAGGAATCTTGGACTTTGACACCTAATGGAT 2400  
QY 2901 ATATTTGAGAGAGAACATCCTTTGGATGTATTTAGATGGGGGACCAAGAGCTG 2960  
Db 2901 ATATTTGAGAGAGAACATCCTTTGGATGTATTTAGATGGGGGACCAAGAGCTG 2960

Db 2401 ATATTTGAGAGAACAGATCACTTTGGATGTAATTAGTTGAGATGGGCAACAAGAGCTG 2460  
QY 2961 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAAGGAGTGCAGAGA 3020  
Db 2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAAGGAGTGCAGAGA 2520  
QY 3021 CTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGA 3080  
Db 2521 CTTATCTCCGGACAACAAGGCTTTAAACCATATTTAACTTTGAACACCTCTGGTAGTGA 2580  
QY 3081 ACAATCTTATAGATCTGCTCCGATGATCAATGAAGATTTCAGTCTGGGAGAGAGATG 3140  
Db 2581 ACAATCTTATAGATCTGCTCCGATGATCAATGAAGATTTCAGTCTGGGAGAGAGATG 2640  
QY 3141 CAAAGTACAGTTTCGAGACACAGACATGGAGTCAATGAGGTGAATCTTCAACAGATAC 3200  
Db 2641 CAAAGTACAGTTTCGAGACACAGACATGGAGTCAATGAGGTGAATCTTCAACAGATAC 2700  
QY 3201 AATATTTCAAGATTTCAGAAAGGTTTGTAAACAGAACTATGGAAAGATACACTCACCGG 3260  
Db 2701 AATATTTCAAGATTTCAGAAAGGTTTGTAAACAGAACTATGGAAAGATACACTCACCGG 2760  
QY 3261 AGAAAGAGATTTCGAGAAACACACACCATGCCAATGAACGAATCTTATTCATGGG 3320  
Db 2761 AGAAAGAGATTTCGAGAAACACACACCATGCCAATGAACGAATCTTATTCATGGG 2820  
QY 3321 TCTCCTTTTGTCAATGCAATATACAAAGGCTTTGTAAAGGCATGCGTACATAGT 3380  
Db 2821 TCTCCTTTTGTCAATGCAATATACAAAGGCTTTGTAAAGGCATGCGTACATAGT 2880  
QY 3381 GGTATGTTGGAGCTGGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATGTA 3440  
Db 2881 GGTATGTTGGAGCTGGCATTTATTTGCTGAAACTCTTCCAAAGCAATCAATATGTA 2940  
QY 3441 TATGGAATTTGGAGGAGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTACATTTGC 3500  
Db 2941 TATGGAATTTGGAGGAGTACTGGGTGTCAGTTCACAAAGACAGATCTTGTACATTTGC 3000  
QY 3501 CACAGCAGCTGCTCTTTTGGCGGTAACCTTGGGAAAGTCTTCTCCAGTTCAGTGCA 3560  
Db 3001 CACAGCAGCTGCTCTTTTGGCGGTAACCTTGGGAAAGTCTTCTCCAGTTCAGTGCA 3060  
QY 3561 ATGAAATGGCAATCTCTCCAGTTCATCACTCAGTCACTGGTAGGCCAGCTTAAT 3620  
Db 3061 ATGAAATGGCAATCTCTCTCCAGTTCATCACTCAGTCACTGGTAGGCCAGCTTAAT 3120  
QY 3621 GGCCTAGCATTTAGCTGAATATGTTATTTACAGAGAGACAGGCTTATCCTGAGTATTA 3680  
Db 3121 GGCCTAGCATTTAGCTGAATATGTTATTTACAGAGAGACAGGCTTATCCTGAGTATTA 3180  
QY 3681 ATTACTTACAGATTATGAGGCTGAAGGTATGGTGGATATAATAGTATTTTAAAGA 3740  
Db 3181 ATTACTTACAGATTATGAGGCTGAAGGTATGGTGGATATAATAGTATTTTAAAGA 3240  
QY 3741 AACTAATTCCTGAACCTTAAATCATCAACAGCAGTGGCCTCTAGCTTTTACTCCTT 3800  
Db 3241 AACTAATTCCTGAACCTTAAATCATCAACAGCAGTGGCCTCTAGCTTTTACTCCTT 3300  
QY 3801 TGCTGAAAAAAA 3813  
Db 3301 TGCTGAAAAAAA 3313

## RESULT 13

AAC66823 standard: cDNA; 4493 BP.

AC AAC66823;

XX

DT 27-FEB-2001 (first entry)

XX Human tankyrase II coding sequence SEQ ID NO: 1.

DE

Human; tankyrase II; telomere length; signal transduction; ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 1..4002

FT /tag= a  
FT /transl\_except= (pos:595..597,aa:Xaa)  
FT /transl\_except= (pos:625..630,aa:XaaXaa)  
FT /transl\_except= (pos:637..639,aa:Xaa)  
FT /transl\_except= (pos:694..696,aa:Xaa)  
FT /transl\_except= (pos:724..726,aa:Xaa)  
FT /transl\_except= (pos:730..732,aa:Xaa)  
FT /transl\_except= (pos:763..765,aa:Xaa)  
FT /transl\_except= (pos:790..792,aa:Xaa)  
FT /transl\_except= (pos:2176..2178,aa:Xaa)  
FT /transl\_except= (pos:2380..2382,aa:Xaa)  
FT /transl\_except= (pos:2398..2400,aa:Xaa)  
FT /transl\_except= (pos:2464..2469,aa:XaaXaa)  
FT /transl\_except= (pos:3163..3165,aa:Xaa)  
FT /note= "Xaa=unknown"  
FT /product= "tankyrase II"  
FT /partial

WO200061813-A1.

19-OCT-2000.

10-APR-2000; 2000WO-US09558.

09-APR-1999; 99US-0128577.

13-APR-1999; 99US-0129123.

(GERO-) GERON CORP.

Morin GB, Funk WD, Piatyszek MA;

WPI: 2000-679503/66.

P-PSDB; AAB27209.

Novel mammalian Tankyrase II polypeptide and the polynucleotide

encoding the polypeptide useful for modulating or maintaining telomere

length, replicative capacity, apoptosis, chromosome packing or gene

expression -

Claim 1; Fig 2; 52pp; English.

The present sequence is a version of the human tankyrase II coding

sequence. Its protein is thought to be involved in signal transduction in

the cell, and to have binding activity for other telomere-associated

proteins. It is possible that it plays a role in the regulation of

telomere length, thus affecting the replicative ability of the cell. The

protein is useful for ribosylating target proteins, for determining

tankyrase II binding activity in a sample, and for modulating telomere

length in a cell.

Query Match 79.6%; Score 3039; DB 21; Length 4493;

Best Local Similarity 88.5%; Pred. No. 0;

Matches 3101; Conservative 1; Mismatches 375; Indels 27; Gaps 2;

QY 337 AACGAGTCAAGAGCTGTGACGCTGAGAAGGTGAACAGCGCGACACGCGGCGAGGA 396

Db 587 AACCATTCNNAGGCTGTGACCCCTGAAAGGTNAACANCCNCAACACGNGGCGAGGA 646

QY 397 AATCCACCCCGCTGCACCTCCCGCAGGTTTGGCGGAAAGACGTAGTTGAATATTGC 456

Db 647 AATCCACCCCGCTGCACCTCCCGCAGGTTTGGCGGAAAGACCTNNNTAAATATTGC 706

QY 457 TTCAGATGGTGCATAATCTCCAAGCACGTGATGGGGCGCTTATTCCTCTTCATATG 516

Db 707 TTCAAATGGTGCATAATNTCCAANCACCTTATAATATGGGGCGCTTATTCCTCTTCATATG 766

QY	517	CATGCTCTTTTGGTCA	TGCTGAAGTAGTCAAT	CTCCTCTTTTGC	GACATGGTGAGAC	CCCCA	576
DB	767	CATGCTCTTTTGGTCA	TGCTGATAAANATCA	ATCTCTTTTGC	GACATGGTGAGAC	CCCCA	826
QY	577	ATGCTCGAGATAATTG	GAAATATATCTCTCT	CCATGAGCTGCAAT	TAAAGGAAGATTG	636	
DB	827	ATGCTCGAGATAATTG	GAAATATATCTCTCT	CCATGAGCTGCAAT	TAAAGGAAGATTG	886	
QY	637	ATGTTTGGCAATTTGT	GCTGTTTACAGCATG	GAGTGAGCCAAACCA	TCCGAAATACAGATG	696	
DB	887	ATGTTTGGCAATTTGT	GCTGTTTACAGCATG	GAGTGAGCCAAACCA	TCCGAAATACAGATG	946	
QY	697	GGACAGCATTTGGATT	TAGCAGATCCATCTG	CCAAAGCAGTGCTT	ACTGCGTAATTAAGA	756	
DB	947	GGACAGCATTTGGATT	TAGCAGATCCATCTG	CCAAAGCAGTGCTT	ACTGCGTAATTAAGA	1006	
QY	757	AAGATGAACCTCTTT	AGAAAGTGCCAGAG	TGGCAATGAAGAAAAA	TATGATGGCTCTACTCA	816	
DB	1007	AAGATGAACCTTTG	AAAGTGCCAGAGTGG	CAATGAAGAAAAA	TATGATGGCTCTACTCA	1066	
QY	817	CACCATTAATGTCAAC	TGCCACGCAAGTGAT	GGCAGAAAGTCAAC	TCCATACATTTGG	876	
DB	1067	CACCATTAATGTCAAC	TGCCACGCAAGTGAT	GGCAGAAAGTCAAC	TCCATACATTTGG	1126	
QY	877	CAGCAGATATAACAG	AGTAAGATTGTACAG	CTGTACTGCAACATGG	AGCTGATGCTCC	936	
DB	1127	CAGCAGATATAACAG	AGTAAGATTGTACAG	CTGTACTGCAACATGG	AGCTGATGCTCC	1186	
QY	937	ATGCTAAAGATAAAG	TGATCTGGTACCAAT	TACACAATGCCCTGT	CTTATATGGTCAATTATG	996	
DB	1187	ATGCTAAAGATAAAG	TGATCTGGTACCAAT	TACACAATGCCCTGT	CTTATATGGTCAATTATG	1246	
QY	997	AAGTAACTGAACCTT	TGGTCAAGCATGGT	CGCTGTGTAAATGCA	ATGGCACTTTGGGCAAT	1056	
DB	1247	AAGTAACTGAACCTT	TGGTCAAGCATGGT	CGCTGTGTAAATGCA	ATGGCACTTTGGGCAAT	1306	
QY	1057	TCACCTCTCTTCATG	AGCAGCTTCTAAGAC	CAGGGTTGAAAGTAT	GTTCCTCTCTCTTAA	1116	
DB	1307	TCACCTCTCTTCATG	AGCAGCTTCTAAGAC	CAGGGTTGAAAGTAT	GTTCCTCTCTCTTAA	1366	
QY	1117	GTATATGGTGCAGACC	CAACACTGCTCAAT	GTGCACAATAAAGTG	CTATAGACTTTGGCTC	1176	
DB	1367	GTATATGGTGCAGACC	CAACACTGCTCAAT	GTGCACAATAAAGTG	CTATAGACTTTGGCTC	1426	
QY	1177	CCACACCAGATTTAA	AGAAAGATTAGCAT	ATGAATTTAAAGGCC	ACCTCGTTGCTGCGAAG	1236	
DB	1427	CCACACCAGATTTAA	AGAAAGATTAGCAT	ATGAATTTAAAGGCC	ACCTCGTTGCTGCGAAG	1486	
QY	1237	CTGCACGAGAGCTG	ATGTTACTCGAATCA	AAAAACATCTCTCT	CTGGAATGGTGAATT	1296	
DB	1487	CTGCACGAGAGCTG	ATGTTACTCGAATCA	AAAAACATCTCTCT	CTGGAATGGTGAATT	1546	
QY	1297	TCAGCATCCTCAAC	CACATGAACAGCAT	TGCCAATTTGCTGCT	GCATCTCCATATCCCA	1356	
DB	1547	TCAGCATCCTCAAC	CACATGAACAGCAT	TGCCAATTTGCTGCT	GCATCTCCATATCCCA	1606	
QY	1357	AAAGAAGCAAAATAT	GTGAACCTTTGCTA	AGAAAAAGGAGCAAA	ACATCAATGAAAAGACTA	1416	
DB	1607	AAAGAAGCAAAATAT	GTGAACCTTTGCTA	AGAAAAAGGAGCAAA	ACATCAATGAAAAGACTA	1666	
QY	1417	AGAATTTCTTGACTC	CTCTGACAGTGGCAT	CTGAGAAAGCTCA	TAAATGATGTTGTTGAAG	1476	
DB	1667	AGAATTTCTTGACTC	CTCTGACAGTGGCAT	CTGAGAAAGCTCA	TAAATGATGTTGTTGAAG	1726	
QY	1477	TAGTGTGAAACATGA	AGCAAAAGGTTAAT	GTCTCTGATTAATCT	TTGGTCAGACTTCTCTAC	1536	
DB	1727	TAGTGTGAAACATGA	AGCAAAAGGTTAAT	GTCTCTGATTAATCT	TTGGTCAGACTTCTCTAC	1786	
QY	1537	ACAGAGCTGCATATT	GTGTCATCTACAAAC	CTGCCGCTACTCCT	CTGAGCTATGGGTGTG	1596	
DB	1787	ACAGAGCTGCATATT	GTGTCATCTACAAAC	CTGCCGCTACTCCT	CTGAGCTATGGGTGTG	1846	

Db 2927 GAAGCCAGGAGGACATGACAGATGCTCTCTCTTCAGGTCATCTAGCCCATCAAGCCTTT 2986  
Qy 2731 CTGAGCAGCAGCTCTTGACAACTTATCTGGAGTGTCTTTCAGAACTGTCTTTCAGTAGTTA 2790  
Db 2987 CTGAGCAGCAGCTCTTGACAACTTATCTGGAGTGTCTTTCAGAACTGTCTTTCAGTAGTTA 3046  
Qy 2791 GTTCAAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCCAGGAGTAGATT 2850  
Db 3047 GTTCAAGTGAACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCCAGGAGTAGATT 3106  
Qy 2851 TTACATAAATCAATTCGTAAAGAACTCTTGACCTTGAGCACCTTAATGGATATATTTTGAGA 2910  
Db 3107 TTACATAAATCAATTCGTAAAGAACTCTTGACCTTGAGCACCTTAATGGATATATTTNAGA 3166  
Qy 2911 GAGAACAGATCACTTTGGATGTATTAAGTTCAGATGGGGCACAGGAGCTGAAGAGATTG 2970  
Db 3167 GAGAACAGATCACTTTGGATGTATTAAGTTCAGATGGGGCACAGGAGCTGAAGAGATTG 3226  
Qy 2971 GAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATCTCCG 3030  
Db 3227 GWATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATCTCCG 3286  
Qy 3031 GACAAAGGCTTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACTAATTTCTTA 3090  
Db 3287 GACAAAGGCTTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACTAATTTCTTA 3346  
Qy 3091 TAGATCTCTCTCGATGATGAAGAGTTTCACTCTGTGGAGGAGAGATGCAAGATACAG 3150  
Db 3347 TAGATCTCTCTCGATGATGAAGAGTTTCACTCTGTGGAGGAGAGATGCAAGATACAG 3406  
Qy 3151 TTCGAGACACAGAGATGGAGTCTAGCTGAGTGAATCTTCAACAGATACAATATTTCTCA 3210  
Db 3407 TTCGAGACACAGAGATGGAGTCTAGCTGAGTGAATCTTCAACAGATACAATATTTCTCA 3466  
Qy 3211 AGATTTCAGAAAGTGTGTAAACAGAACTATG- - - - -GGAAAGAT 3249  
Db 3467 AGATTTCAGAAAGTGTGTAAACANNNNNNNNNNNGAGCCAGATTTCGACGAGCAAGAT 3526  
Qy 3250 ACATCACCAGGAGAAAGATTTCTGAAGAAAACACCAACCATGCCAATGAACGAATGC 3309  
Db 3527 ACATCACCAGGAGAAAGATTTCTGAAGAAAACCAACCAATGCCAATGAACGAATGC 3586  
Qy 3310 TATTTTCATGGTCTCTCTTTTGGAGTGGCAATATATCCACAAAGGCTTTGATGAAAGGCATG 3369  
Db 3587 TATTTTCATGGTCTCTCTTTTGGAGTGGCAATATATCCACAAAGGCTTTGATGAAAGGCATG 3646  
Qy 3370 CGTACATAGTGGTATGTTTGGAGTGGCAATATATTTTGTGAAACTCTTTCCAAAAGCA 3429  
Db 3647 CGTACATAGTGGTATGTTTGGAGTGGCAATATATTTTGTGAAACTCTTTCCAAAAGCA 3706  
Qy 3430 ATCAATATGTATATGGAATTCGAGAGAGTACTGGGTGTCAGATTTCACAAAGCAGATCTT 3489  
Db 3707 ATCAATATGTATATGGAATTCGAGAGAGTACTGGGTGTCAGATTTCACAAAGCAGATCTT 3766  
Qy 3490 GTTACATTTGCCACAGGAGCTGCTCTTTTGGCGGTGAACCTTTGGGAAAGTCTTTCTTCG 3549  
Db 3767 GTTACATTTGCCACAGGAGCTGCTCTTTTGGCGGTGAACCTTTGGGAAAGTCTTTCTTCG 3826  
Qy 3550 AGTTCAGTCAATGAAATGGACATTCCTCCAGGTCATCATCTACTGCTAGGTCAGGC 3609  
Db 3827 AGTTCAGTCAATGAAATGGACATTCCTCCAGGTCATCATCTACTGCTAGGTCAGGC 3886  
Qy 3610 CCAGTGAATAGCCCTAGCATAGCTGAATGTTTATTTACAGAGGAAACAGGCTTATC 3669  
Db 3887 CCAGTGAATAGCCCTAGCATAGCTGAATGTTTATTTACAGAGGAAACAGGCTTATC 3946  
Qy 3670 CTGAGTATTAATTAATCTTACCAGATATGAGGCCTCAAGGATGTTGTCGATGATAAATAG 3729  
Db 3947 CTGAGTATTAATTAATCTTACCAGATATGAGGCCTCAAGGATGTTGTCGATGATAAATAG 4006  
Qy 3730 TTAATTTAAGAACTAATTTCCACTGAACCTAAATCATCAAGAGCAGTGGCCTCTACG 3789

Db 4007 TTATTTTAAAGAACTAATTTCCACTGAACCTAAATCATCAAGCAGCAGTGGCCTCTACG 4066  
Qy 3790 TTTTACTCTTTGCTGAAAAAAA 3813  
Db 4067 TTTTACTCTTTGCTGAAAAAAA 4090  
RESULT 14  
AAC89706  
ID AAC89706 standard; cDNA: 5005 BP.  
XX AAC89706;  
AC  
XX 08-MAR-2001 (first entry)  
DT  
XX Human adult T-cell leukaemia cDNA MO-BC-203.  
DE  
XX Human; cytostatic; gene therapy; vaccine; breast cancer;  
KW T cell leukaemia; cancer associated antigen; ss.  
OS Homo sapiens.  
XX  
XX WO200072021-A2.  
PN  
XX 30-NOV-2000.  
XX  
XX 24-MAY-2000; 2000WO-US14391.  
PF  
XX 26-MAY-1999; 99US-0320092.  
PR  
XX (LUDW-) LUDWIG INST CANCER RES.  
PA  
XX Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;  
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;  
PI Sahlin U, Pfreundschuh M, Old LJ, Knuth A, Jager E;  
XX WPI: 2001-032067/04.  
DR P-PSDB: AAB48574.  
XX  
PT Diagnosing disorder characterized by expression of human cancer  
PT associated antigen precursor, involves detecting interaction of an  
PT agent with novel NA Group 1 nucleic acid molecule encoding the antigen  
PT precursor  
XX  
PS Example 8; Page 82-84; 87pp; English.  
CC The present sequence is given in a specification relating to nucleic  
CC acids and encoded polypeptides which are cancer associated antigens  
CC expressed in patients afflicted with breast cancer and/or T cell  
CC leukemia. A disorder characterised by expression of human cancer  
CC associated antigen may be diagnosed by contacting a biological sample  
CC with an agent that specifically binds to the nucleic acid that encodes  
CC the antigen, complexed with human leukocyte antigen molecule, and  
CC determining the interaction between the agent and the nucleic acid.  
CC Cancer associated antigens, the nucleotides encoding them, antibodies  
CC against them and the pharmaceutical compositions comprising them are  
CC useful for diagnosing, monitoring and treating the diseases characterised  
CC by the expression of one or more cancer associated antigens.  
XX  
SQ Sequence 5005 BP; 1492 A; 917 C; 1015 G; 1571 T; 10 other;  
Query Match 72.7%; Score 2774.6; DB 22; Length 5005;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2810; Conservative 0; Mismatches 4; Indels 3; Gaps 3;  
Qy 997 AAGTAACCTGAACTTTTGGTCAAGCATGTCCTCTGTAATGCAATGGACTTTGTGGCAAT 1056  
Db 1 AAGTAACCTGAACTTTTGGTCAAGCATGTCCTCTGTAATGCAATGGACTTTGTGGCAAT 60  
Qy 1057 TCACCTCTCTTCATGAGCAGCTTCTTAAGACAGGGTTGAAGTATGTCTCTCTCTTAA 1116  
Db 61 TCACCTCTCTTCATGAGCAGCTTCTTAAGACAGGGTTGAAGTATGTCTCTCTCTTAA 120

QY 1117 GTTATGTCAGAGCCCAACACTGCTCAATTTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176  
Db 121 GTTATGTCAGAGCCCAACACTGCTCAATTTGTCACAATAAAAGTGCTATAGACTTGGCTC 180  
QY 1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTCGAAG 1236  
Db 181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTCGAAG 240  
QY 1237 CTGCACGAGAGCTGTATGTTACTCGAATCAAAACATCTCTCTCGAAATGGTGAAT 1296  
Db 241 CTGCACGAGAGCTGTATGTTACTCGAATCAAAACATCTCTCTCGAAATGGTGAAT 300  
QY 1297 TCACAGCATCTCAACACATGAACACAGCATGTGCTGTGCTGTGATCTCCATATCCCA 1356  
Db 301 TCACAGCATCTCAACACATGAACACAGCATGTGCTGTGCTGTGATCTCCATATCCCA 360  
QY 1357 AAAGAAACCAATATGTGAACCTGTGCTAAGAAAAGAGCAACATCAATGAAGACAT 1416  
Db 361 AAAGAAACCAATATGTGAACCTGTGCTAAGAAAAGAGCAACATCAATGAAGACAT 420  
QY 1417 AAGAAATCTTCTGACTCTCTGCTGCACTGGCATCTGAGAAAGCTCATAATGATGTTGAAG 1476  
Db 421 AAGAAATCTTCTGACTCTCTGCTGCACTGGCATCTGAGAAAGCTCATAATGATGTTGAAG 480  
QY 1477 TAGTGTGAAACATGAAGCAAGGTTAATGCTCTGCTGATTAATCTTGTCAGACTTCTCTAC 1536  
Db 481 TAGTGTGAAACATGAAGCAAGGTTAATGCTCTGCTGATTAATCTTGTCAGACTTCTCTAC 540  
QY 1537 ACAGAGCTGCATATGTTGGTTCATCTACAAACCTGCGCCCTACTCTGAGCTATGGTGTG 1596  
Db 541 ACAGAGCTGCATATGTTGGTTCATCTACAAACCTGCGCCCTACTCTGAGCTATGGTGTG 600  
QY 1597 ATCCTTAACATTTATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTAC 1656  
Db 601 ATCCTTAACATTTATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAATGTAC 660  
QY 1657 AGCAACTCTCCAAAGAGGTATCTCATTAAGGTAATTCAGAGGCACACAGCAATTCGTCGG 1716  
Db 661 AGCAACTCTCCAAAGAGGTATCTCATTAAGGTAATTCAGAGGCACACAGCAATTCGTCGG 720  
QY 1717 AAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACTGTGTACTGTTCAGAGTGCA 1776  
Db 721 AAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAAAACTGTGTACTGTTCAGAGTGCA 780  
QY 1777 ACTGCAGACACATGAAGGCGCTCAGTCTACACACTTCATTTTCAGCTGGGTATAACA 1836  
Db 781 ACTGCAGACACATGAAGGCGCTCAGTCTACACACTTCATTTTCAGCTGGGTATAACA 840  
QY 1837 GAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGTGCTGATGCTAAAGATAAG 1896  
Db 841 GAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGTGCTGATGCTAAAGATAAG 900  
QY 1897 GAGGCTCTTACCTTTGCAACATGCAATGCTTTATGAGCATTTACAGTTCGAGAACTTC 1956  
Db 901 GAGGCTCTTACCTTTGCAACATGCAATGCTTTATGAGCATTTACAGTTCGAGAACTTC 960  
QY 1957 TTGTTAAACATGAGCAGTAGTATGAGCTGATTTATGAAATTTACACCTTTACATG 2016  
Db 961 TTGTTAAACATGAGCAGTAGTATGAGCTGATTTATGAAATTTACACCTTTACATG 1020  
QY 2017 AAGCAGCAGCAAAAGGAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTCAGACC 2076  
Db 1021 AAGCAGCAGCAAAAGGAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTCAGACC 1080  
QY 2077 CTACCAAAAAAACAGGATGGAATATCTCCTTTGGATCTTTGTTAAAGATGGAGATACAG 2136  
Db 1081 CTACCAAAAAAACAGGATGGAATATCTCCTTTGGATCTTTGTTAAAGATGGAGATACAG 1140  
QY 2137 ATATTCAAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTT 2196  
Db 1141 ATATTCAAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTT 1200  
QY 2197 TAGCCAGAGTGAAGAAGTTGCTTCTCTGATATGTAATGTAATGCGCGGATACCCCAAGCA 2256

Db 1201 TAGCCAGAGTGAAGAAGTTGCTTCTCTGATATATGTAATGCGCGGATACCCCAAGCA 1260  
QY 2257 GACATTCAACACCTTTACATTTAGCAGCTGGTTATTAATAATTTAGAAGTTGCAGAGTATT 2316  
Db 1261 GACATTCAACACCTTTACATTTAGCAGCTGGTTATTAATAATTTAGAAGTTGCAGAGTATT 1320  
QY 2317 TGTTTACAACACGGAGCTGATGTGAATGCCCAGACAAAGGAGGACTTATTCCTTTACATA 2376  
Db 1321 TGTTTACAACACGGAGCTGATGTGAATGCCCAGACAAAGGAGGACTTATTCCTTTACATA 1380  
QY 2377 ATGCAGCATCTTACGGGCAITAGATGTAGCAGCTCTACTAATAAAGTATATCATGTG 2436  
Db 1381 ATGCAGCATCTTACGGGCAITAGATGTAGCAGCTCTACTAATAAAGTATATCATGTG 1440  
QY 2437 TCAATGCCACGACCAATATGGGCTTTACACCTTTGCAGAGCAGCCCAAAAGGACGAA 2496  
Db 1441 TCAATGCCACGACCAATATGGGCTTTACACCTTTGCAGAGCAGCCCAAAAGGACGAA 1500  
QY 2497 CACAGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 2556  
Db 1501 CACAGCTTTGTGCTTTGTGTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 1560  
QY 2557 GACAAACACCTTTAGATTTAGTTTACAGCGATGATGTACAGCGCTCTTCTGACAGCAGCCA 2616  
Db 1561 GACAAACACCTTTAGATTTAGTTTACAGCGATGATGTACAGCGCTCTTCTGACAGCAGCCA 1620  
QY 2617 TGCCCCCATCTGCTCTGCCCTTTGTTACAAAGCCTCAAGTGTCTCAATGGTGTGAGAGCC 2676  
Db 1621 TGCCCCCATCTGCTCTGCCCTTTGTTACAAAGCCTCAAGTGTCTCAATGGTGTGAGAGCC 1680  
QY 2677 CAGGAGCCACTGACAGATGCTCTCTCTCAGGTCCATCTAGCCCATCAAGCCTTCTTCGAG 2736  
Db 1681 CAGGAGCCACTGACAGATGCTCTCTCAGGTCCATCTAGCCCATCAAGCCTTCTTCGAG 1739  
QY 2737 CCAGAGCTTTGACAACCTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 2796  
Db 1740 CCAGAGCTTTGACAACCTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 1797  
QY 2797 GTGGAACAGAGGGTGTCTCCAGTTTGAGAAAAGAGAGTTCCAGGAGTAGATTTAGCA 2856  
Db 1798 GTGGAACAGAGGGTGTCTCCAGTTTGAGAAAAGAGAGTTCCAGGAGTAGATTTAGCA 1857  
QY 2857 TAACCTAAATTCGTAAGGAATCTTGGACTTGAGCACCTTAATGGATATATTTGAGAGAGAAC 2916  
Db 1858 TAACCTAAATTCGTAAGGAATCTTGGACTTGAGCACCTTAATGGATATATTTGAGAGAGAAC 1917  
QY 2917 AGATCATTCTGGATGTATAGTTGAGATGGGACACAGAGCTGAAGGAGATTGGAATCA 2976  
Db 1918 AGATCATTCTGGATGTATAGTTGAGATGGGACACAGAGCTGAAGGAGATTGGAATCA 1977  
QY 2977 ATGCTTATGGACATAGGCACAACTAATTAAGAGAGTCCAGAGACTTATCTCCGGACAAC 3036  
Db 1978 ATGCTTATGGACATAGGCACAACTAATTAAGAGAGTCCAGAGACTTATCTCCGGACAAC 2037  
QY 3037 AAGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGAACAATTTCTTATAGATC 3096  
Db 2038 AAGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGAACAATTTCTTATAGATC 2097  
QY 3097 TGTCTCTGTATGAAGAGTTTCACTCTGTGGAGAGAGATGCAAGTTCAGAGTCGAG 3156  
Db 2098 TGTCTCTGTATGAAGAGTTTCACTCTGTGGAGAGAGATGCAAGTTCAGAGTCGAG 2157  
QY 3157 AGCAGAGATGGAGTCTATGAGTGGATCTTCAACAGATACATATTTCTCAAGATTC 3216  
Db 2158 AGCAGAGATGGAGTCTATGAGTGGATCTTCAACAGATACATATTTCTCAAGATTC 2217  
QY 3217 AGAAGTTTGTAAACAAAGAACTATGGAAAGATACACTCACCGGAGAAAAGAGTTCTG 3276  
Db 2218 AGAAGTTTGTAAACAAAGAACTATGGAAAGATACACTCACCGGAGAAAAGAGTTCTG 2277  
QY 3277 AAGAAACACACACCATGCCAATGAACGAATGCTATTTTCATGGGCTCTCTTTTGTGAATG 3336

Db 2278 AAGAAACCAACACCATGCGCAATGAACGATGCTATTTCATGGGTCTCCTTTTGTGAATG 2337  
 Qy 3337 CAATATCCACAAAGGCTTTGATGAAGAGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3396  
 Db 2338 CAATATCCACAAAGGCTTTGATGAAGAGCATGCGTACATAGGTGGTATGTTTGGAGCTG 2397  
 Qy 3397 GCATTTATTTGGCTGAAACCTTTCCAAAGCAATCAATATGATGGAATGGAGGAG 3456  
 Db 2398 GCATTTATTTGGCTGAAACCTTTCCAAAGCAATCAATATGATGGAATGGAGGAG 2457  
 Qy 3457 GTACTGGGTGCCAGTTTACAAAGACAGATCTTGTACATTTGCCACAGGAGCTGCTCT 3516  
 Db 2458 GTACTGGGTGCCAGTTTACAAAGACAGATCTTGTACATTTGCCACAGGAGCTGCTCT 2517  
 Qy 3517 TTTGCCGGGTAACTTGGGAAAGTCTTTCCHGCAAGTTTCAGTGCATGCAATGAAATGGCAATT 3576  
 Db 2518 TTTGCCGGGTAACTTGGGAAAGTCTTTCCTGCAGTTTCAGTGCATGCAATGAAATGGCAATT 2577  
 Qy 3577 CTCTCCAGGTCACTCACPCAGTCACTGGTAGGCCAGTGTAAATGGGCTAGCATTAGCTG 3636  
 Db 2578 CTCTCCAGGTCACTCACPCAGTCACTGGTAGGCCAGTGTAAATGGGCTAGCATTAGCTG 2637  
 Qy 3637 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTTACCAGATTA 3696  
 Db 2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTTACCAGATTA 2697  
 Qy 3697 TGAGGCCCTGAAGGTATGTCGATGATGAATAAGTATTTTAAAGAACTAATTTCCACTGAA 3756  
 Db 2698 TGAGGCCCTGAAGGTATGTCGATGATGAATAAGTATTTTAAAGAACTAATTTCCACTGAA 2757  
 Qy 3757 CCTAAATCATCAAGAGCAGTGGCCCTCTAGCTTTTACTCCTTTGCTGAAAAAAA 3813  
 Db 2758 CCTAAATCATCAAGAGCAGTGGCCCTCTAGCTTTTACTCCTTTGCTGAAAAAAA 2814

RESULT 15  
 AAF63919  
 ID AAF63919 standard; DNA; 2971 BP.  
 XX AAF63919;  
 XX  
 DT 05-APR-2001 (first entry)  
 XX  
 DE Human tankyrase2 clone consensus SEQ ID NO: 92.  
 XX  
 KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
 KW Inflammatory disorder; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200100849-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 28-JUN-2000; 2000WO-US17827.  
 XX  
 PR 29-JUN-1999; 99US-0141582.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 XX  
 DR WPI; 2001-102896/11.  
 XX  
 PT New tankyrase2 polypeptides, useful for treating conditions mediated by  
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
 PT inflammatory and autoimmune disorders -  
 XX  
 PS Example 1; Page 152-153; 242pp; English.  
 XX  
 CC The present invention provides the protein and coding sequence for the  
 CC human tankyrase2 protein. This is found in two different versions,  
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has

CC polyADP-ribosylation activity and is involved in the modification of  
 CC TRF1, which is a telomere-specific binding protein. The regulation of  
 CC telomere length, in which TRF1 has a role, is linked to aging and  
 CC cancer. The sequences are useful in the treatment of cancers and  
 CC inflammatory disorders.

XX  
 SQ Sequence 2971 BP; 915 A; 555 C; 642 G; 859 T; 0 other;

Query Match 63.9%; Score 2439.8; DB 22; Length 2971;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1371 TGTGAACCTGTTGCTAAGAAAAGAGCAACATCAATGAAAGACTAAAGAAATTCCTGACT 1430  
 Db 1 TGTGAACCTGTTGCTAAGAAAAGAGCAACATCAATGAAAGACTAAAGAAATTCCTGACT 60  
 Qy 1431 CCTCTGCACGTGCATCTGAGAAGCTCATAATGATGTTGTTGAAAGTAGTGGTGAACAT 1490  
 Db 61 CCTCTGCACGTGCATCTGAGAAGCTCATAATGATGTTGTTGAAAGTAGTGGTGAACAT 120  
 Qy 1491 GAAGCAAAGGTTAATGCTCTGGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550  
 Db 121 GAAGCAAAGGTTAATGCTCTGGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180  
 Qy 1551 TGTGGTCACTACAAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTGATCCTTAACATATA 1610  
 Db 181 TGTGGTCACTACAAAACCTGCCGCCCTACTCCTGAGCTATGGGTGTGATCCTTAACATATA 240  
 Qy 1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAA 1670  
 Db 241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAATGAAATGTACAGCAACTCCTCCAA 300  
 Qy 1671 GAGGGTATCTCATTAGGTAATTCAGAGGACAGACAGCAATTCGCTGGAAGCTGCAAGAGCT 1730  
 Db 301 GAGGGTATCTCATTAGGTAATTCAGAGGACAGACAGCAATTCGCTGGAAGCTGCAAGAGCT 360  
 Qy 1731 GGAGATGTCGAAACTGTAAACAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACTT 1790  
 Db 361 GGAGATGTCGAAACTGTAAACAACTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACTT 420  
 Qy 1791 GAAGGGCTGAGTCTACACCACCTCATTTGCGAGCTGGGTATACAGAGTGCCTGGTG 1850  
 Db 421 GAAGGGCTGAGTCTACACCACCTCATTTGCGAGCTGGGTATACAGAGTGCCTGGTG 480  
 Qy 1851 GAATATCTGCTACAGCATGGAGCTGATGCTGCTAAAGATAAAGGAGGCTTGTACCT 1910  
 Db 481 GAATATCTGCTACAGCATGGAGCTGATGCTGCTAAAGATAAAGGAGGCTTGTACCT 540  
 Qy 1911 TTGCACAATGCATGCTTCTATGGACATTAAGTTCGCAACTTCTTTTAAACATGGA 1970  
 Db 541 TTGCACAATGCATGCTTCTATGGACATTAAGTTCGCAACTTCTTTTAAACATGGA 600  
 Qy 1971 GCAGTATTAATGCTAGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAA 2030  
 Db 601 GCAGTATTAATGCTAGCTGATTTATGAAATTTTACACCTTTTACATGAAGCAGCAGCAAAA 660  
 Qy 2031 GGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCACAAAAAAC 2090  
 Db 661 GGAAATATGAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACCACAAAAAAC 720  
 Qy 2091 AGGGATGAAATACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATCAAGATCTG 2150  
 Db 721 AGGGATGAAATACTCCTTTGGATCTTTGTTAAAGATGGAGATACAGATATCAAGATCTG 780  
 Qy 2151 CTTAGGGAGATGAGCTTTTCTAGATGCTGCCAAGAGGGTCTTTAGCCAGAGTGAAG 2210  
 Db 781 CTTAGGGAGATGAGCTTTTCTAGATGCTGCCAAGAGGGTCTTTAGCCAGAGTGAAG 840  
 Qy 2211 AAGTTGCTCTTCTCTGATAATGTAATTTGCCGCGATACCCCAAGGACAGACATTCACACCT 2270  
 Db 841 AAGTTGCTCTTCTCTGATAATGTAATTTGCCGCGATACCCCAAGGACAGACATTCACACCT 900  
 Qy 2271 TTACATTTAGCAGCTGGTGTATAAATTTAGAAAGTTGACAGAGTATTTGTTTACAACACGGA 2330

Db 901 TTAGATTAGCAGCTGGTTATAATAATTTAGAAGCTTGCAGAGTATTTGTTTACAACCGGA 960  
QY 2331 GCTGATGTGAATGCCAAGACAAAGAGGAGCTTATTCCTTTTACATAATGACAGCATCTTAC 2330  
Db 961 GCTGATGTGAATGCCAAGACAAAGAGGAGCTTATTCCTTTTACATAATGACAGCATCTTAC 1020  
QY 2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATTAATGATGTGTCAATGCCACCGAC 2450  
Db 1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATTAATGATGTGTCAATGCCACCGAC 1080  
QY 2451 AAATGGGCTTTTCAACACCTTTTGCAGGAGCAGCCOCAAAGGGAGCAACACAGCTTTTGCT 2510  
Db 1081 AAATGGGCTTTTCAACACCTTTTGCAGGAGCAGCCOCAAAGGGAGCAACACAGCTTTTGCT 1140  
QY 2511 TTGTTGCTAGGCCATGAGCTGACCCGACCTTAAANAATCAGGAAGACAAACACCTTTTA 2570  
Db 1141 TTGTTGCTAGGCCATGAGCTGACCCGACCTTAAANAATCAGGAAGCAACACCTTTTA 1200  
QY 2571 GATTTAGTTTACGGGATGATGTACGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630  
Db 1201 GATTTAGTTTACGGGATGATGTACGGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260  
QY 2631 CTGCCCTCTTGTACAGCCTCAAGTCTCAATGGTGTGAGAGCCOCCAGGAGCCOACATGCA 2690  
Db 1261 CTGCCCTCTTGTACAGCCTCAAGTCTCAATGGTGTGAGAGCCOCCAGGAGCCOACATGCA 1320  
QY 2691 GATGCTCTCTCTCAGCTCCATCTAGCCCATCAAGCCCTTCTGCAGCCAGCAGCTCTTGAC 2750  
Db 1321 GATGCTCTCTCTCAGCTCCATCTAGCCCATCAAGCCCTTCTGCAGCCAGCAGCTCTTGAC 1380  
QY 2751 AACTTATCTGGGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGCT 2810  
Db 1381 AACTTATCTGGGAGTTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGCT 1440  
QY 2811 GCTTCCAGTTTGGAGAAAAGGAGTTCCAGAGTACATTTTAGCATTAACATCAATTCGTA 2870  
Db 1441 GCTTCCAGTTTGGAGAAAAGGAGTTCCAGAGTACATTTTAGCATTAACATCAATTCGTA 1500  
QY 2871 AGGAATCTTGGACTTGGACACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930  
Db 1501 AGGAATCTTGGACTTGGACACCTTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560  
QY 2931 GTATTAGTTGAGATGGGACACAGGAGCTGAAGAGATGGAATCAATGCTTATGGACAT 2990  
Db 1561 GTATTAGTTGAGATGGGACACAGGAGCTGAAGAGATGGAATCAATGCTTATGGACAT 1620  
QY 2991 AGGCACAACTAATTAAGAGTTCGAGAGACTTATCTCCGGACAAACAAGTCTTTAACCCA 3050  
Db 1621 AGGCACAACTAATTAAGAGTTCGAGAGACTTATCTCCGGACAAACAAGTCTTTAACCCA 1680  
QY 3051 TATTTAACTTTGAACACCTCTGGTAGTGAACAATTTCTTATAGATCTGCTCCTGATGAT 3110  
Db 1681 TATTTAACTTTGAACACCTCTGGTAGTGAACAATTTCTTATAGATCTGCTCCTGATGAT 1740  
QY 3111 AAAGAGCTTTTCACTCTGAGGAGAGATGCAAAAGTACAGTTGAGAGACAGAGATGGA 3170  
Db 1741 AAAGAGCTTTTCACTCTGAGGAGAGATGCAAAAGTACAGTTGAGAGACAGAGATGGA 1800  
QY 3171 GGTATGCAAGTGGAACTCTTCAACAGATCAATATTTCTCAAGATTCAGAAGGTTTGTAAAC 3230  
Db 1801 GGTATGCAAGTGGAACTCTTCAACAGATCAATATTTCTCAAGATTCAGAAGGTTTGTAAAC 1860  
QY 3231 AAGAACTATGGGAAGATACACTACCGGAGAAAGAAAGTTTCTGAAGAAACCAACAC 3290  
Db 1861 AAGAACTATGGGAAGATACACTACCGGAGAAAGAAAGTTTCTGAAGAAACCAACAC 1920  
QY 3291 CATGCCAATGAAGATGCTATTTCATGGTCTCCTTTTGTGAATGCAATTTATCCACAA 3350  
Db 1921 CATGCCAATGAAGATGCTATTTCATGGTCTCCTTTTGTGAATGCAATTTATCCACAA 1980  
QY 3351 GGCTTTGATGAAGGCATGCGTACATAGGTGCTATCTTTGGAGCTGGCATTTTATTTGCT 3410

Db 1981 GGCTTTGATGAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGTCT 2040  
QY 3411 GAAAACTCTTCCAAAAGCAATCAATATATATGAAATTTGGAGGAGTACTGGGTGTCCA 3470  
Db 2041 GAAAACTCTTCCAAAAGCAATCAATATATGAAATTTGGAGGAGTACTGGGTGTCCA 2100  
QY 3471 GTTCCACAAAGACAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCCGGTAACC 3530  
Db 2101 GTTCCACAAAGACAGATCTTGTACATTTGCCACAGGAGCTGCTCTTTTGGCCGGTAACC 2160  
QY 3531 TTGGGAAAGTCTTTCCCTGCGAGTTTCAGTGCATGAAATGGCACATTTCTCTCCAGGTCAT 3590  
Db 2161 TTGGGAAAGTCTTTCCCTGCGAGTTTCAGTGCATGAAATGGCACATTTCTCTCCAGGTCAT 2220  
QY 3591 CACTCAGTCACTGCTAGGCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTTATTAC 3650  
Db 2221 CACTCAGTCACTGCTAGGCCAGTGTAAATGGCCCTAGCATTTAGCTGAATATGTTTATTAC 2280  
QY 3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTAATTTACTTACCAGATTTATGAGCCCTGAAGGT 3710  
Db 2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTAATTTACTTACCAGATTTATGAGCCCTGAAGGT 2340  
QY 3711 ATGGTCGATGGATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAATATCAATCAA 3770  
Db 2341 ATGGTCGATGGATAAATAGTATTTTAAAGAACTAATTTCCACTGAACCTAAATATCAATCAA 2400  
QY 3771 AGCAGCAGTGGCCTCTACGTTTACTCCTTTGCTGAAAAAAA 3813  
Db 2401 AGCAGCAGTGGCCTCTACGTTTACTCCTTTGCTGAAAAAAA 2443

Search completed: February 11, 2003, 13:10:38  
Job time : 752.794 secs





|||||  
Db 248 TATTTGGTTTCAGAAATGTCACAGACGTGATGATGGGGCCTTATTCCTCTT 307  
QY 510 CATAATGCATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGGACATGGTGCA 569  
Db 308 CATAAATGCATGCTCTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGGACATGGTGCA 367  
QY 570 GACCCCAATGCTCGAGATAATGGAATATATACCTCTCCATGAAGCTGCAATTAAGGA 629  
Db 368 GACCCCAATGCTCGAGATAATGGAATATATACCTCTCCATGAAGCTGCAATTAAGGA 427  
QY 630 AAGATTGATGTTGCAATGCTGCTTTACAGATGGAGCTGAGCCAAACCTCCGAAATACA 689  
Db 428 AAGATTGATGTTGCAATGCTGCTTTACAGATGGAGCTGAGCCAAACCTCCGAAATACA 487  
QY 690 GATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTACTGGTGA 749  
Db 488 GATGGAAGGACAGCATTTGGATTTAGCAGATCCATCTGCCAAAGCAGTGTACTGGTGA 547  
QY 750 TATAAGAAAGATGAATCTTTAGAAAGTGCAGGAGTGGCAATGAAGAAATAATGATGGCT 809  
Db 548 TATAAGAAAGATGAATCTTTAGAAAGTGCAGGAGTGGCAATGAAGAAATAATGATGGCT 607  
QY 810 CTACTCACACATTAATGTCACAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 869  
Db 608 CTACTCACACATTAATGTCACAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 667  
QY 870 CATTTGGCAGCAGGATATAACAGATTAAGATTGTACAGCTGTTACTGCAACATGGAGCT 929  
Db 668 CATTTGGCAGCAGGATATAACAGATTAAGATTGTACAGCTGTTACTGCAACATGGAGCT 727  
QY 930 GATGTCATGCTTAAGATAAAGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 989  
Db 728 GATGTCATGCTTAAGATAAAGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 787  
QY 990 CATTATGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGGAGCTTG 1049  
Db 788 CATTATGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGTAAATGCAATGGAGCTTG 847  
QY 1050 TGGCAATTAACCTCTTTCATAGGAGCTTCTAAGAACAGGGTTGAAAGTATGTTCTCTT 1109  
Db 848 TGGCAATTAACCTCTTTCATAGGAGCTTCTAAGAACAGGGTTGAAAGTATGTTCTCTT 907  
QY 1110 CTCTTAAGTATGCTGCAGACCCAAACACTGCTCAATGTGCACAAATGAAGTGTATAGAC 1169  
Db 908 CTCTTAAGTATGCTGCAGACCCAAACACTGCTCAATGTGCACAAATGAAGTGTATAGAC 967  
QY 1170 TTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229  
Db 968 TTGGCTCCACACACAGTTAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027  
QY 1230 CTGCAAGCTGCAGGAGAGCTGATGTTACTCGAATCAAAACATCTCTCTGGAATG 1289  
Db 1028 CTGCAAGCTGCAGGAGAGCTGATGTTACTCGAATCAAAACATCTCTCTGGAATG 1087  
QY 1290 GTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCAATGCTGCTGCAATCTCA 1349  
Db 1088 GTGAATTTCAAGCATCTCAACACATGAACAGCATTTGCAATGCTGCTGCAATCTCA 1147  
QY 1350 TATCCCAAGAAAGCAATATGTGAACCTGTTGCTAAGAAAGGAGCAACATCAATGAA 1409  
Db 1148 TATCCCAAGAAAGCAATATGTGAACCTGTTGCTAAGAAAGGAGCAACATCAATGAA 1207  
QY 1410 AAGACTAAAGAAATTTTGACTCTCTGACGTGGCATCTGAGAAAGCTCATATGATGTT 1469  
Db 1208 AAGACTAAAGAAATTTTGACTCTCTGACGTGGCATCTGAGAAAGCTCATATGATGTT 1267  
QY 1470 GTTGAAGTAGTGTGAACATGAACAAAGGTTAATGCTCTGATAATCTTGGTCAGACT 1529  
Db 1268 GTTGAAGTAGTGTGAACATGAACAAAGGTTAATGCTCTGATAATCTTGGTCAGACT 1327  
QY 1530 TCTCTACACAGAGCTGCATTTGTTGCTATCTACAAACCTGCGGCTACTCTGAGCTAT 1589  
|||||

Db 1328 TCTCTACACAGAGCTGCATATTTGGTCACTCAAAACCTGCCGCTACTCTCTGAGCTAT 1387  
QY 1590 GGGTGTGATCCTTAACATATATCCCTTCAGGGCTTTTACTGCTTTACAGATGGGAAATGAA 1649  
Db 1388 GGGTGTGATCCTTAACATATATCCCTTCAGGGCTTTTACTGCTTTACAGATGGGAAATGAA 1447  
QY 1650 AATGTACAGCAACCTCCCAAGAGGGTATCTCATTAGTAAATTCAGAGGCGACAGACAA 1709  
Db 1448 AATGTACAGCAACCTCCCAAGAGGGTATCTCATTAGTAAATTCAGAGGCGACAGACAA 1507  
QY 1710 TTGCTGGAAGCTGCAAAAGGCTGGAGATGTGAAACTGTAAAAAACTGTGTACTGTTGAG 1769  
Db 1508 TTGCTGGAAGCTGCAAAAGGCTGGAGATGTGAAACTGTAAAAAACTGTGTACTGTTGAG 1567  
QY 1770 AGTCTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCACCTTCAATTTGACAGCTGGG 1829  
Db 1568 AGTCTCAACTGCAGAGACATTTGAAGGGCGTCAGTCTACACCACCTTCAATTTGACAGCTGGG 1627  
QY 1830 TATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1889  
Db 1628 TATAACAGAGTGTCCGTGGTGGAAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1687  
QY 1890 GATAAGGAGGCTTGPACCTTTGCAATGCAATGTTCTTATGGACATTTATGAAGTTGCA 1949  
Db 1688 GATAAGGRRGSCCTTGTACCTTTGCAATGCAATGTTCTTATGGACATTTATGAAGTTGCA 1747  
QY 1950 GAACTTCTTCTTAAACATGAGCAGTAGTTAATGTAGTCTGATTTATGGAATTTTACACCT 2009  
Db 1748 GAACTTCTTCTTAAACATGAGCAGTAGTTAATGTAGTCTGATTTATGGAATTTTACACCT 1807  
QY 2010 TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069  
Db 1808 TTACATGAAGCAGCAGCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 1867  
QY 2070 GCAGACCTTACCAAAAAACACAGGATGGAATACTCCTTTGGATCTTGTAAAGATGGA 2129  
Db 1868 GCAGACCTTACCAAAAAACACAGGATGGAATACTCCTTTGGATCTTGTAAAGATGGA 1927  
QY 2130 GATACAGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAG 2189  
Db 1928 GATACAGATATTCAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAG 1987  
QY 2190 GGTGTTTAGCCAGAGTGAAGAGTTGCTCTCTCTGATATGTAATGTAATGCTCCGATACC 2249  
Db 1988 AGTGTGTTAGCCAGAGTGAAGAGTTGCTCTCTCTGATATGTAATGTAATGCTCCGATACC 2047  
QY 2250 CAAGGCGACATTCACACCTTTACATTTAGCAGCTGGTGTATATATTTAGAGATTGCA 2309  
Db 2048 CAAGGCGACATTCACACCTTTACATTTAGCAGCTGGTGTATATATTTAGAGATTGCA 2107  
QY 2310 GAGTATTTGTTACACACGAGCTGATGTAATGCCCAAGCAAAAGGAGGACTTATTCCT 2369  
Db 2108 GAGTATTTGTTACACACGAGCTGATGTAATGCCCAAGCAAAAGGAGGACTTATTCCT 2167  
QY 2370 TTACATAATGTCAGCATCTTACGGGCATGATGTTAGCAGCTCTACTAAATTAAGTATAAT 2429  
Db 2168 TTACATAATGTCAGCATCTTACGGGCATGATGTTAGCAGCTCTACTAAATTAAGTATAAT 2227  
QY 2430 GCATGTGTCATTCACGAGCAATGCGGCTTTACACCTTTGACCAAGAGGAGGAGGAGGAG 2489  
Db 2228 GCATGTGTCATTCACGAGCAATGCGGCTTTACACCTTTGACCAAGAGGAGGAGGAGGAG 2287  
QY 2490 GGACCAACACAGCTTTGCTGCTTTGCTAGCCATGGAGCTGACCCGACTCTTAAAT 2549  
Db 2288 GGACCAACACAGCTTTGCTGCTTTGCTAGCCATGGAGCTGACCCGACTCTTAAAT 2347  
QY 2550 CAGGAAGGACAAACACCTTTAGATTAGTTTACGGGATGATGTACGCGCTCTTCTGACA 2609  
Db 2348 CAGGAAGGACAAACACCTTTAGATTAGTTTACGGGATGATGTACGCGCTCTTCTGACA 2407  
QY 2610 GCAGCATGCCCCCATCTGCTGCTGCTGCTTGTACAGCCCTCAAGTGTCAATGTTG 2669  
Db 2408 GCAGCATGCCCCCATCTGCTGCTGCTGCTTGTGTACAGCCCTCAAGTGTCAATGTTG 2467

QY 2670 AGAAGCCAGGAGGACCTGAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729  
DB 2468 AGAAGCCAGGAGGACCTGAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2527  
QY 2730 TCTGCAGCCAGCAGCTTGAACACTTATCTGGGAGTTTTCAGAACTGCTCTTCAGTAGTT 2789  
DB 2528 TCTGCAGCCAGGAGCTTGAACACTTATCTGGGAGTTTTCAGAACTGCTCTTCAGTAGTT 2587  
QY 2790 AGTTCAAGTGGAAACAGAGGCTTCCAGTTTGGAGAAAAAGGAGTTTCCAGGAGTAGAT 2849  
DB 2588 AGTTCAAGTGGAAACAGAGGCTTCCAGTTTGGAGAAAAAGGAGTTTCCAGGAGTAGAT 2647  
QY 2850 TTTAGCAATACTCAATTCGTAGGAATCTTGGACTTGGACACTTAATGGATATATTTCAG 2909  
DB 2648 TTTAGCAATACTCAATTCGTAGGAATCTTGGACTTGGACACTTAATGGATATATTTCAG 2707  
QY 2910 AGAGAACAGATCACCTTTGGATGATTAAGTGTAGATGGGGCACAGGAGCTGAAGAGAGATT 2969  
DB 2708 AGAGAACAGATCACCTTTGGATGATTAAGTGTAGATGGGGCACAGGAGCTGAAGAGAGATT 2767  
QY 2970 GGAATCAATGCTTTATGGACATAGGCACAACTTAATTAAGGAGTGGAGAGACTTATCTCC 3029  
DB 2768 GGAATCAATGCTTTATGGACATAGGCACAACTTAATTAAGGAGTGGAGAGACTTATCTCC 2827  
QY 3030 GGACAAACAGGCTTTAAACCCATTTAACTTTGAACACTCTGGTAGTGGAACTTCTT 3089  
DB 2828 GGACAAACAGGCTTTAAACCCATTTAACTTTGAACACTCTGGTAGTGGAACTTCTT 2887  
QY 3090 ATAGATCTCTCTCTCATATAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAAAGTACA 3149  
DB 2888 ATAGATCTCTCTCTCATATAAGAGTTTCACTCTGTGGAGGAAGAGATGCAAAAGTACA 2947  
QY 3150 GTTCGAGACACAGAGATGGAGGTCATGAGGTGGAATCTTCAACAGATACAAATTTCTC 3209  
DB 2948 GTTCGAGACACAGAGATGGAGGTCATGAGGTGGAATCTTCAACAGATACAAATTTCTC 3007  
QY 3210 AGATTTCAAGGTTTGTAAACAGAACTATGGAAGAGATACACACCGGAGAAAGAA 3269  
DB 3008 AGATTTCAAGGTTTGTAAACAGAACTATGGAAGAGATACACACCGGAGAAAGAA 3067  
QY 3270 GTTCTCTGAAGAAACACACACCATGCAATGAACGAATGCTATTTTCATGGGTCTCTTTT 3329  
DB 3068 GTTCTCTGAAGAAACACACACCATGCAATGAACGAATGCTATTTTCATGGGTCTCTTTT 3127  
QY 3330 GTGAATGCAATTTATCCAAAGAGCTTTGTATGAAGGATCGCTACATAGGTGGTAGTTT 3389  
DB 3128 GTGAATGCAATTTATCCAAAGAGCTTTGTATGAAGGATCGCTACATAGGTGGTAGTTT 3187  
QY 3390 GGAGCTGGCATTTATTTTCTGCTGAAACTCTTCCAAAGCAATCAATATATATGAAATT 3449  
DB 3188 GGAGCTGGCATTTATTTTCTGCTGAAACTCTTCCAAAGCAATCAATATATGAAATT 3247  
QY 3450 GGAGGAGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTACATTTGCCACAGGAC 3509  
DB 3248 GGAGGAGGTACTGGGTGTCAGTTTCAAAAGACAGATCTTGTACATTTGCCACAGGAC 3307  
QY 3510 CTGCTCTTTTGGCGGTAACTTTGGGAAAGTCTTTCCTGCAGTTCAGTCAATGAAAGT 3569  
DB 3308 CTGCTCTTTTGGCGGTAACTTTGGGAAAGTCTTTCCTGCAGTTCAGTCAATGAAAGT 3367  
QY 3570 GCACATTTCTCTCCAGGTCTATCACTCAGTCTGCTAGGCGCCAGTGAATGGCCTTAGCA 3629  
DB 3368 GCACATTTCTCTCCAGGTCTATCACTCAGTCTGCTAGGCGCCAGTGAATGGCCTTAGCA 3427  
QY 3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCTCTGAGTATTTAATTACTTAC 3689  
DB 3428 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCTCTGAGTATTTAATTACTTAC 3487  
QY 3690 CAGATTATGAGGCTGAGGATGCTGATGATAAATAGTTATTTTAAAGAACTAATTC 3749  
DB 3488 CAGATTATGAGGCTGAGGATGCTGATGATAAATAGTTATTTTAAAGAACTAATTC 3547

QY 3750 CACTGAACCTAAATCATCAAGCAGCAGTGGCTCTACGTTTACTCTCTTTGCTGAAAA 3809  
DB 3548 CACTGAACCTAAATCATCAAGCAGCAGTGGCTCTACGTTTACTCTCTTTGCTGAAAA 3607  
QY 3810 AAAA 3813  
DB 3608 AAAA 3611  
RESULT 2  
US-09-350-982C-4  
; Sequence 4, Application US/09350982C  
; Patent No. 6455290  
; GENERAL INFORMATION:  
; APPLICANT: Berthelsen, Jens  
; APPLICANT: Toma, Salvatore  
; APPLICANT: Isacchi, Antonella  
; TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: PHRM-0043  
; CURRENT APPLICATION NUMBER: US/09/350,982C  
; CURRENT FILING DATE: 1999-07-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 3498  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: PCR Primers  
; NAME/KEY: CDS  
; LOCATION: (1)..(3498)  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc\_feature  
; LOCATION: (1102)..(1102)  
; OTHER INFORMATION: n is any nucleic acid  
; NAME/KEY: misc\_feature  
; LOCATION: (2650)..(2650)  
; OTHER INFORMATION: n is any nucleic acid  
US-09-350-982C-4  
Query Match 91.4%; Score 3488.8; DB 4; Length 3498;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 3487; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
QY 225 ATGTCGGTCCCGCTGCGCGGGGAGCGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCC 284  
DB 1 ATGTCGGTCCCGCTGCGCGGGGAGCGCTGCGGAGCGCTGCGGAGCGCGCGCGCGCGCGCC 60  
QY 285 GTGGAGCGCGCGCGCGCGCGCGCTGTCGAGGCGTCCCAACGGGAGCGTGAACGAGTC 344  
DB 61 GTGGAGCGCGCGCGCGCGCGCGCTGTCGAGGCGTCCCAACGGGAGCGTGAACGAGTC 120  
QY 345 AAGAGCTGTGACCGCTGGAAGGTGAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 404  
DB 121 AAGAGCTGTGACCGCTGGAAGGTGAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 180  
QY 405 CGGCTGCACTTCGCGCGAGGTTTGGCGGAAAGACGTAGTTGAATATTGCTTCAGAAAT 464  
DB 181 CGGCTGCACTTCGCGCGAGGTTTGGCGGAAAGACGTAGTTGAATATTGCTTCAGAAAT 240  
QY 465 GGTGCAAAATGTCAAAGCAGTATGATGGGGCGCTTATTCCTCTTCAATAATGCTGCTCT 524  
DB 241 GGTGCAAAATGTCAAAGCAGTATGATGGGGCGCTTATTCCTCTTCAATAATGCTGCTCT 300  
QY 525 TTTGTCATGCTGAAGTGTGAATCTCTTTTGGACATGTTGCGAGACCGCGCGCGCGCGCGCG 584  
DB 301 TTTGTCATGCTGAAGTGTGAATCTCTTTTGGACATGTTGCGAGACCGCGCGCGCGCGCGCG 360  
QY 585 GATAATTGGAATTATATCTCTCCATGAAGCTGCAATTAAGGAAAGATTTGATTTTTCG 644  
DB 361 GATAATTGGAATTATATCTCTCTCTCCATGAAGCTGCAATTAAGGAAAGATTTGATTTTTCG 420



Db	2581		GAGGGTCTTCACGTTTGAGAAAAAGAGGTTCCAGAGTAGAATTTAGCAATACCTCAA	2640
Qy	2865		TTCGTAAGGAATCTTGGACCTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACT	2924
Db	2641		TTCGTAAGGNATCTTGGACTTGAGCACCTTAATGGATATATTTGAGAGAGAACAGATCACT	2700
Qy	2925		TTGGATGTATATTAGTTGAGATGGGGCACAAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT	2984
Db	2701		TTGGATGCTATTAGTTGAGATGGGGCACAAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT	2760
Qy	2985		GGACATAGGCACAACTAATTAAAGAGCTCGAGAGACTTATCTCCGCACAAACAGGCTTT	3044
Db	2761		GGACATAGGCACAACTAATTAAAGAGCTCGAGAGACTTATCTCCGCACAAACAGGCTTT	2820
Qy	3045		AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACTTCTTATAGATCTGCTCCCT	3100
Db	2821		AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACTTCTTATAGATCTGCTCCCT	2880
Qy	3105		GATGATAAGAGGTTTCAGTCTCTGGAGGAAGAGATGCAAGTACAGTTTCAGAGACACAGA	3164
Db	2881		GATGATAAGAGGTTTCAGTCTCTGGAGGAAGAGATGCAAGTACAGTTTCAGAGACACAGA	2940
Qy	3165		GATGGAGGTCATCGAGTGGAACTTTCAACAGATACAAATATTCFCAAGATTTCAGAAGGTT	3224
Db	2941		GATGGAGGTCATCGAGTGGAACTTTCAACAGATACAAATATTCFCAAGATTTCAGAAGGTT	3000
Qy	3225		TGTAACAAGAAACTATGGGAAAAGATACACTACCCGGAGAAAAGGTTTCTGAAGAAAAC	3284
Db	3001		TGTAACAAGAAACTATGGGAAAAGATACACTACCCGGAGAAAAGGTTTCTGAAGAAAAC	3060
Qy	3285		CACAACTATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTGTGAATGCAATTATC	3344
Db	3061		CACAACTATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTGTGAATGCAATTATC	3120
Qy	3345		CACAAAGGCTTTGATCAAAAGGCATGCGTACATAGTGGTATGTTTGGAGCTGGCATTTAT	3404
Db	3121		CACAAAGGCTTTGATCAAAAGGCATGCGTACATAGTGGTATGTTTGGAGCTGGCATTTAT	3180
Qy	3405		TTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGATATGGAATTGGAGGAGGTACGTGG	3464
Db	3181		TTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGATATGGAATTGGAGGAGGTACGTGG	3240
Qy	3465		TGTCCAGTTTCACAAACAGACATCTGTTACATTTTGCACAGGACGCTGCTCTTTTGCCGG	3524
Db	3241		TGTCCAGTTTCACAAACAGACATCTGTTACATTTTGCACAGGACGCTGCTCTTTTGCCGG	3300
Qy	3525		GTAACTTTGGGAAAGCTTTTCTCGAGTTTCAGTGCAAATGAAAATGGCACATTTCTCTCCA	3584
Db	3301		GTAACTTTGGGAAAGCTTTTCTCGAGTTTCAGTGCAAATGAAAATGGCACATTTCTCTCCA	3360
Qy	3585		GGTCATCACTCAGTCACTGGTAGGCCAGTGTAATGGCTTAGCATTTAGCTGGAATATGTT	3644
Db	3361		GGTCATCACTCAGTCACTGGTAGGCCAGTGTAATGGCTTAGCATTTAGCTGGAATATGTT	3420
Qy	3645		ATTTTACAGAGGAGAACAGGCTTATCTGAGTATTTTAAATTACTTACCAGATTATGAGGCTT	3704
Db	3421		ATTTTACAGAGGAGAACAGGCTTATCTGAGTATTTTAAATTACTTACCAGATTATGAGGCTT	3480
Qy	3705		GAAGGTATGGTCGATGGA	3722
Db	3481		GAAGGTATGGTCGATGGA	3498

### RESULT 3

```

US-09-196-387-1
;
; Sequence 1, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; OF USE THEREOF
;

```

```

: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue, 4th Floor
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/196,387
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/095,225
: FILING DATE: June 10, 1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-230 CIPI
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-487-5800
: TELEFAX: 201-343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4134 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: US-09-196-387-1

Query Match          49.4%; Score 1885; DB 4; Length 4134;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps

Qy 252 GGAGCGGCTGCGGACGCGCGCGGCGGAGCGCGTGAGCGCGCGCGCGAGAGCTGTC 311
Db 507 GGGCGACGAGGACCTGGGACAGGGGTCCCAGCAGTGCAGCGGGGCCCTACGGGAACCTGCTG 566
Qy 312 GAGCGGTGCGCAACGGGGACGTGGAACAGTCAAGAGGCTGTGACGCGCTGAGAAGGTG 371
Db 567 GAGCGCTGTGCAATGGGACGTGTCCGGGTAAAGAGGCTGTGGACGGCGCAACGTA 626
Qy 372 AACAGCGCGACGCGGCGGAGAAATCCACCCGCTGCACATTGCGCGAGGTTTTGGG 431
Db 627 AATCAAAAGGACATGGCGCGCGGAAGCTTTCTCCCTGCACATTCGCTGCAGGTTTTTGA 686
Qy 432 CGGAAGACGCTAGTTGAATATTGCTTCAGAATGGTSCAAATGTCCAAGCACGCTGATGAT 491
Db 687 AGGAAGATGTTGAGAACACTTACTACAGATGGTGCTAATGTCCACCGCTCGTGATGAT 746
Qy 492 GGGGGCCTTATTCTCTTCATAATGATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db 747 GGAGGTCTCAFPCCGCTTCATAATGCGCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806
Qy 552 CTTTTCGACATGGTCAGACCCCCAATGCTCGAGATAAATTGGAAATTATCTCCTCTCCAT 611
Db 807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866
Qy 612 GAAGCTGCAATTAAAGAAAGATTGATGTTTGCATTGCTGTTTACAGCATGGAGCTGAG 671
Db 867 GAAGCTGCTAATAAGGAAGATCGATGTGTGCATTGTGCTGTGCAGCACGCGAGCTGAC 926
Qy 672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTTGATTTAGCAGATTCATCTGCCAA 731
Db 927 CCAACATCTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCTCTCAGCAAA 986

```

Qy	732	GCAGTGCCTTACTGGTGAATATATAGAAAGATGAACCTCTTTAGAAAGTGCACAGAGTGGCAAT	731
Db	987	CGTGTCCCTTACAGGTAATACAAAGAAAGACGAACCTCTAGAAAGTGTGTAGAGTGGTAAT	1046
Qy	792	GAAGAAAAATGATGCTCTACTCACACATTAAATGTCAACTGCCAGCGAAGTGATGGC	851
Db	1047	GAGAAAAATTAATGGCTTTACTGACTCTCTAAATGTGAATTGGCATGCAAGTGTATGGG	1106
Qy	852	AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGATAAGATTTGTACAGCTG	911
Db	1107	CGAAAGTCGACTCCTTTACATCTAGCAGGGCTACACAGAGTTCGAATAGTTCAGCTT	1166
Qy	912	TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATTAAGGTGTATCTGGTACCAATTCAC	971
Db	1167	CTTCTTCAGCATGGTCTGATGTTATGCAAAAGACAAAGGTGGACTTGTGCCCTTCTCAT	1226
Qy	972	AATGCCCTGTCTTATGCTCATATGAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTGT	1031
Db	1227	ATGTCATGTTTCATATGGACATTTATGAAGTCAGACAGTGTCTACTTAAAGCATGGAGCTTGT	1286
Qy	1032	GTAATGCAATGGACTTTGTGCAATTTCACTCTCTCTTCATGAGGCGAGCTTCTTAAGAACAGG	1091
Db	1287	GTTAATGCCATGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCCT	1346
Qy	1092	GTTGAAGTATGTTCTCTCTCTTTAAGTTTATGGTCAGACCCCAACACTGCTCTCAATGTCAC	1151
Db	1347	GTAGAAGTCTGCTTTTGTGTACTTATGCCATGGCGCTGATCCTTAGCTTAGTCAACTGCCAT	1406
Qy	1152	AATAAAAAGTCTATACACTTTGCTCTCCACACACAGTTTAAAGAAAGATTAGCATATGAA	1211
Db	1407	GGCAAAAGTCTGTGGATATGCTCCAATCCGGAGCTTTGGGAGAGATGACTTTATGAA	1466
Qy	1212	TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA	1271
Db	1467	TTTAAAGTCAATCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAGATTAAAAA	1526
Qy	1272	CATCTCTCTCGAAATGGTGAATTTCAAGCATCTCAACACATGAACACAGATTCGAT	1331
Db	1527	ACACTGCTCTGGAATTCATTAATTTCAACAAACCGCAGTCTCATGAACACAGCACTGCAC	1586
Qy	1332	TGTGCTGCTCATCTCCATATCCCAAAAGAAAGCAAAATATGTGAACCTGTTGCTAAGAAAA	1391
Db	1587	TGTGCTGTGCCCTCTCGCATCCCAACGTAACCAAGTGCAGAAATTTGTACTTTAGAAAA	1646
Qy	1392	GGAGCAACATCAATGAAAGACTAAAGAAATTTCTTGACTCTCTGACGCTGGCATCTGAG	1451
Db	1647	GGAGCAAAATGTAATGAAAAAATAAAGATTTTCATGACTCCCTTGCATGTTGCAGCGCAA	1706
Qy	1452	AAAGCTCATATCATGTTGTTGAGTGTGGTGAACATGAAGCAAAAGGTTAATGCTCTG	1511
Db	1707	AGAGCCCATTAATGATGTCATGGAAGTTCTGCAATAAGCATGGCGCAAGATGAATGCACTG	1766
Qy	1512	GATAATCTTGGTCAGACTTCTCTACACAGCTGCATATTTGTGTCTATCTACAAACCTGC	1571
Db	1767	GACACCTTGGTCAGACTGCTTTGCATAGAGCGGCCCTAGCAGGCCACCTGCAGACCTGC	1826
Qy	1572	CGCCTACTCTGAGCTATGGGTGTGATGCTAAACATTATATCCCTTCAGGGCTTTACTGCT	1631
Db	1827	CGCCTCTGCTGAGTTACGGCTCTCACCCCTCCATCATCTCTCTTACAAGGCTTCACAGCA	1886
Qy	1632	TTACAGATGGAAATCAAAATCTACAGCAACTCTCTCCAAGAGGGTATCTCATTAGGTAAT	1691
Db	1887	GCACAGATGGCAATGAAGCAGTGCAGCAGATTTCTGAGTGAGATGACACCTTATACGTAAT	1946
Qy	1692	TCAGAGCAGACAGACAATTTGCTGGAAGCTGCAAAAGGCTGGAGATGTCGAAACTGTAAAA	1751
Db	1947	TCGTGATGTGATTATCGACTCTTAGAGGCATCTAAGCTGGAGACTTGGAACTGTGAAG	2006
Qy	1752	AAACTGTGTACTCTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTGTACTTACACCA	1811
Db	2007	CAACTTTGAGCTCTCAAAATGTGAATTTAGAGACTTTAGAGGCGCGGCATTTCCACGCC	2066

Qy	1812	CTTCAITTTGCACCTGGGTATACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGA	187
Db	2067	TTACACTTCGCACAGGCTACACCGCGTCTGTGTGTAGAGTACTGCTACACACGGT	2126
Qy	1872	GCTGATGTGATCTAAAGATAAAGAGGCGCTTGTAACCTTTGCACAATGCATGTTCTTAT	1931
Db	2127	GCCGATGTCATGCCAAGACACAGGTGGCTTGGTGCCCTTCATAATGCCCTGTTTCAT	2186
Qy	1932	GGACATTTAAGATTTCGAAACTTCTTGTTAACATGGAGCAGTAGTTATGTAGCTCAT	1991
Db	2187	GGACACTATGAGTGGCTGAGCTTTTAGTAAGCATGGGCTTCTGTCAATGTGGCGAC	2246
Qy	1992	TTATGGAAATTTACACCTTTACATGAACAGACAGCAAAAGAAATATGAATTTGCAAA	2051
Db	2247	TTATGGAAATTTACCCTCTCCATGAAGCAGCAGCTAAAGAAAGTATGAATCTGCAAG	2306
Qy	2052	CTTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAAACACAGGATGGAATACTCCTTG	2111
Db	2307	CTCCTTTTAAACATGAGCAGATCCAACCTAAAAAGAACAGAGATGGAATACACCTTG	2366
Qy	2112	GATCTGTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAGGGGAGATGCAGCTTG	2171
Db	2367	GATTTGGTAAAGAGGAGACACAGATATTCAGACTTTACTTAAAGGGGATGCTGCTTG	2426
Qy	2172	CTAGATCTGCCAAGAAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTCTTCCTCATAT	2231
Db	2427	TTGGATGCTCCRAAGAGGCTGCCGACAGAGTGCAGAGCTGTACCCACAGAAAT	2486
Qy	2232	GTAAATFGCCGCTATCCCAAGGCAGACAATCAACACCTTTACATTTAGCAGCTGGTTAT	2291
Db	2487	ATCAACTGCAGACACCCAGGCGAATAATCCACCCCTCTGCACCTGGCAGCAGGCTAT	2546
Qy	2292	AATAATTTAGAAGTTCACAGATATTGTTTACAAACAGCAGCTGATGTAATGCCCAAGAC	2351
Db	2547	AATACTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGAC	2606
Qy	2352	AAAGGAGGACTTATTCCTTTTACATATGACAGATCTTACGGGCATGTAGATGTAGCAGCT	2411
Db	2607	AAGSGTGTTTAAATTCCTTTCATATGCGCATCTTATGGCATGTTGCATAGCGGCT	2666
Qy	2412	CTACTATAAAGTATAATCATGTGTCAATGCCAGGACAAATGGCTTTCACACCTTG	2471
Db	2667	TTATTGATAAATACAAACAGCTGTGTAATGAACAGATAAGTGGCGCTTTACTCCCTC	2726
Qy	2472	CAGCAGCAGCCCAAAAGGAGCAACACAGCTTTGTGCTTTGTGTAGCCCATGGAGCT	2531
Db	2727	CATGAAGCAGCCAGAAAGGAAGCAGCAGCTGCGCCCTCTCTCTAGCAGCATGGTCA	2786
Qy	2532	GACCCGACTCTTAAATCAGGAAGCAAAACACCTTTAGATTTAGTTTACAGCGGATGAT	2591
Db	2787	GACCCACCATGAAGAACAGGAAGGCCAGACGCCCTCGATCTGGCAACAGCTGAGCAT	2846
Qy	2592	GTACAGGCTCTTCTGACAGCAGCATGCCCCCATCTGCTGCCCCCTTGTTTACAAGCTT	2651
Db	2847	ATCAGACTTTGCTGATAGATGCCATGCCCTCCAGAGGCTTTACCTACCTGTTTAAACCT	2906
Qy	2552	CAAGTGTCTAATGGTGTGAGAGCCACAGGACCACTGCAGATGCTCTCTCTCAGGTCCA	2711
Db	2907	CAGG-----CTACTGTAGTGAAGTGCGCTCTCTGATCTCACCAGCA	2945
Qy	2712	TCATGCCCATCAAGCGCTTCTGCAGCCAGCAGCTTGACAACCTATCTGGGAGTTTTCA	2771
Db	2946	TCCACCCCTCTCGCTCTCGGCTGCCAGCAGATAGACAACCTCAGTGGCCCTTTAGCA	3005
Qy	2772	GAACGTGCTTACGTAGTTAGTTCAAGTGGAAACAGAGGTCCTCCAGTTTGGAGAAAAG	2831
Db	3006	GAGTTGGCCGTAGGAGGAGCCCTCCAATGCAGGGGATGCGCGGGGAACAGAAAGGAAG	3065
Qy	2832	-----GAGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTGAAGGAATCTTGGACTT	2885
Db	3066	GAGGAGAAGTGTGCTGGTCTTGACATGAATATCAGCCCAATTTCTAAAAGGCTTGGCGCT	3125
Qy	2886	GAGCACTTAATGATATATTTTGGAGAGAAACAGATCACTTTGGATGTATATTAGTTGAGATG	2945











QY 2399 AGATGTAGCAGCTCTACTAATAAAGTATATAATGCATGTCTCAATGCCAGGACAAATGGCC 2458  
DB 2967 TGACATAGCGGCTTATTCATATAATACACACAGCTGTCTAATGCAACAGATAAGTGGCC 3026  
QY 2459 TTTCACACCTTTGCGAGAGCAGCCCAAGAGGAGACACAGCTTTGTGCTTTGCTTGTCT 2518  
DB 3027 GTTTACTCCCTCCATGAAGCAGCCAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3086  
QY 2519 AGCCCATGAGCTGAGCCGACCTCTTAAATAATCAGGAAGGACAAACACCTTTAGATTTAGT 2578  
DB 3087 AGCGATGTGACAGCCCAACCATGAAGAACAGGAAGCCAGAGCCCTCTGGATCTGGC 3146  
QY 2579 TTCCAGGGATGATGTGAGCGCTCTCTCT----- 2605  
DB 3147 AACAGCTGACGATATCAGAGCTTTGACATGTTGACATAGCGGCTTTATGATAAATAACAA 3206  
QY 2606 ----- 2605  
DB 3207 CAGCTGTGTAATGCAACAGATAAGTGGCGGTTTACTCCCTCCATGAAGCAGCCAGAA 3266  
QY 2606 -----GACAGCAGCCATGCCCATCTCTCTGCTCTGCTCTGTTTACAAGCCTCAAG----- 2655  
DB 3267 AGGAAGACGACGCTGTGCGCCCTCTCTAGCGATGTGCGACCCACCCATGAAGAA 3326  
QY 2656 -----TGCTCAA 2662  
DB 3327 CCAGGAAGCCACAGCCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGAT 3386  
QY 2663 TGTGTGAGAGCCCGAGGAGCCACTGC----- 2689  
DB 3387 AGATGCCATGCCCCAGAGGCGCTTACCTACTGTTTAAACCTCAGGCTACTGTAGTGAG 3446  
QY 2690 AGATGCTCTCTCTCAGTCTCAGTCCATCAGCCATCAGCCCTTTCTGAGCCAGCAGCTTGA 2749  
DB 3447 TGCCCTCTGATCTCACCAGCATCCACCCTCTCTGCTCTGCTGCTGCGCTGCGCAGCAGATAGA 3506  
QY 2750 CAATCTATCTGGGAGTCTTTCAGAACTGTCTTCAGTAGTGTAGTTCAAGTGGAAACAGAGG 2809  
DB 3507 CAACCTCAGTGGCCCTTAGCAGAGTTGGCCGTAGGAGGAGCTCCAATGCGAGGATGG 3566  
QY 2810 TGCTTCAGCTTTGGAGAAAGAG-----GTTCCAGAGGTAGATTTTAGCATAACTCA 2863  
DB 3567 CGCCGGGGAAACAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3626  
QY 2864 ATTCTGAAGCAATCTGGAGCTTGCAGCCTTAATGGATATATTGAGAGAGACAGATCAC 2923  
DB 3627 ATTCTAAAGAGCTTGGCCCTTGAACACCTTCGGGATATCTTTGAACAGAGACAGATTAC 3686  
QY 2924 TTTGGATGATTATTAGTTGAGATGGGACAGAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2983  
DB 3687 ACTAGATGTTGCTGATGATGGTCTATGAAGAGTTGAAGAAATAGGCATCAATGCATA 3746  
QY 2984 TGGACATAGGCAACAACATAATTAAGAGGTGAGAGCTTATGCTCCGGACACAAAGTCT 3043  
DB 3747 TGGGCGCCGCCAACAAATTAATCAAGAGGTAGAAAGACTTTAGGTGGACACAGAGGCAC 3806  
QY 3044 TAACCCATATTAACTTTGAACACCTCTGCTAGTGAACAAATTTATAGATCTGCTCC 3103  
DB 3807 CAATCCCTTATTGACTTTTTCACCTGCTGTTAATCAGGAAAGGATTTTCTGCTGCTCTCC 3866  
QY 3104 TGATGATAAAGAGTTTTCAGTCTCTGGAGGAGAGATGCAAGATPACAGTTTCGAGAGCACAG 3163  
DB 3867 AGAATAAAGATATACAGTCTAGTGGAGAGAGATGCAAGTACTATTTCGAGAACACAG 3926  
QY 3164 AGATGGAGGTATGACAGGTGGAATCTTCAACAGATACATATTTCTCAAGATTGAGAGGT 3223  
DB 3927 AGATGGTGGTAAATGCTGGCGGCTCTTCAACAGATACAAATGTCTCGAATTCAAAAAGT 3986  
QY 3224 TTGTAAACAAGAACTATGGGAAGATACACTACCCGAGAGAAAGAGTTTCTGAGAAAAA 3283  
DB 3987 TGTAACAAGAAAGTTGAGGAGGGTCTTCCACCAGAGAGGAAAGTGTCTGAGAGAA 4046

QY 3284 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3343  
DB 4047 TCACAACCATCACAAATGAGCGCATGTTGTTTCATGGTCTCCTTTTCATTAATGCAATTAT 4106  
QY 3344 CCACAAGGCTTTTCATGAAAGGCGATGCTACATAGTGGTATGTTTGGAGTGGCATTTA 3403  
DB 4107 TCATAAAGGTTTTCATGAGCGACATGCTACATAGGAGGAATGTTTGGGCGCGGATTTA 4166  
QY 3404 TTTTGTCTGAAACCTCTTCCAAAAGCAATCAATATGATATGGAATTTGGAGGAGTACTGG 3463  
DB 4167 TTTTGTCTGAAACCTCTTCCAAAAGCAACCAATATGTTATGGAATTTGGAGGAGCAAGG 4226  
QY 3464 GTGTCCAGTTTCCAAAAGCAGATCTTGTACATTTGGCAGAGCAGCTGCTCTTTTGGCG 3523  
DB 4227 CTGCGCTACACACAGGACAGCTCATGTATATATGATATGATGATGATGATGATGATGAT 4286  
QY 3524 GGTAACTTGGGAAGTCTTCTCTGAGTTCAGTGCAGTCAATGAAATGCAATTTCTCTCC 3583  
DB 4287 AGTGACCTTGGGAATCTTCTCTGAGTTCAGTGCAGTCAATGAAATGCAATTTCTCTCC 4346  
QY 3584 AGGTCTACCTCAGTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3643  
DB 4347 AGGCAACCTCAGTCTGATGTTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4406  
QY 3644 TATTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3703  
DB 4407 CATCTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4466  
QY 3704 TGAAG 3708  
DB 4467 AGAAG 4471

RESULT 6  
US-09-196-387-12  
; Sequence 12, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196,387  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/095,225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 base pairs



QY 2402 TGTAGCAGCTCTACTAATAAAGTATATGATGTGTCATGCGCACGACAAATGGGCTTT 2461  
Db 1182 AATTGTGGAAGTGTACTGAAATACCAAGCTGCGCAATCGAAGCAACTACTGAATCCGGTCT 1241  
QY 2462 CACACCTTTTGCACGAAGCAGCCCAAGGAGGACGACGAGCTTTGTCTTTGTGTAGC 2521  
Db 1242 CTCACCGCTGATGCTGCTTTTATGGGTCATTAACATTTGTCATCTATTACTACA 1301  
QY 2522 CCATGGAGCTGACCGGACTCTTAAATATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581  
Db 1302 ACAAGGTGCTAATGAGATGTGGTACAGTACGGGTCGAAAGCCCTCTCATTTAGCTGC 1361  
QY 2582 A 2582  
Db 1362 A 1362

## RESULT 8

US-09-031-485-36/c  
; Sequence 36 Application US/09031485  
; Patent No. 5824306  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Liang  
; APPLICANT: Blehm, E. Scot  
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; CITY: 1825 Sharp Point Drive  
; STATE: Fort Collins  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: WordPerfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/031.485  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/847.429  
; FILING DATE: 24-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verser, Carol Talkington  
; REGISTRATION NUMBER: 37,459  
; REFERENCE/DOCKET NUMBER: HW-5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 970/493-7272  
; TELEFAX: 970/484-9505  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5235 nucleotides  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-031-485-36

Query Match 2.4%; Score 93; DB 1; Length 5235;  
Best Local Similarity 45.0%; Pred. No. 6.6e-14;  
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAGAACTAAAGAAATTCCTGACTCTCTCGACGTGGCATCTCAG 1451  
Db 5058 GGCACCGATATCAACACATGCAATGCGCTTAATGCAATTCATCTCGCCCTCCAAA 4999

QY 1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAACATATGAAGCAAGGTTAATGCTCTG 1511  
Db 4998 GAAGTTCATCATGAGTGGTCCCGAACTTCTGAAAAGAAAGCAGATCTTGATGCTGCC 4939  
QY 1512 GATAATCTTTGGTCAGACTTCTCTACACAGAGTGCATATTTGGTGCATCTCAAAACCTGC 1571  
Db 4938 ACTAGAAAGGGTAAACACAGCGTTACATATAGCATCATTTGGCAGGACAAGAACTAATCGTC 4879  
QY 1572 CGCTACTCTCGAGCTATGGGTGATCTAAGCATATATATCTCCCTTCAGGCGCTTTACTGCT 1631  
Db 4878 ACAGTACTTTGTTGAAATGGTCTAATGTTAGTGAACATCACTAAACGGTTTTACACCA 4819  
QY 1632 TTACAGATGGGAAATGAAATGTACAGCAACTCTCCAAAGAGGGTATCTCATTAGGTAAT 1691  
Db 4818 CTTTACATGCTGCACAAGAAATCACGAATCTGTTGTAGCTATCTTCTTGCACCAAT 4759  
QY 1692 TCAGAGGCACAGACAATTTGCTGGAGCTGCAAAGGCTGGAGATGT-----CGAAAC 1744  
Db 4758 GCCAATCAAGCTTTAAGTACAGAAAGACGGTTTTACGCCACTGGCAGTTGCCCTTGCACAA 4699  
QY 1745 TGTAAAAAAGCTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCGTCAGTC 1804  
Db 4698 GGTACAGATCGTGTGGTCTGTTTGTGTTGAAATGACACGCGCGGAAAGTG-CGCTT 4640  
QY 1805 TACACCACTTCAATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGTCTACA 1864  
Db 4639 GCCAGCACTGCATATTTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580  
QY 1865 GCATGAGCTGATGTCATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCACAAATGCATG 1924  
Db 4579 AAATGAGCATAACTCGGATGTGACTTTGAAAAGCGCTTTACTCCGCTTCATATCCCGCG 4520  
QY 1925 TTCTTATGGACATTATCAAGTTCAGAACTTCTTGTAAACATGGAGCAGTAGTTAATGT 1984  
Db 4519 TCACATGGAATGAGAACGTACGACAACTGCTACTCGAAGAGGAGCCAAATGTAATTA 4460  
QY 1985 AGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAATATGAAAT 2044  
Db 4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGTCGTACAAACAT 4400  
QY 2045 TTGCAAACTTCTGTCAGCATGGTGCAGACCCCTACCACAAAAGAGGATGGAATAC 2104  
Db 4399 GGTTCGTTATTTGTTGGCTCATGGGCGCCGTAATTGACTGTCGCACACGTAATTTACTAAC 4340  
QY 2105 TCCTTTGGATCTTGTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161  
Db 4339 ACATTTACACTGCTCTCTGTCAGTCATGATCAAGTTGTTGATTTGTTGCTTGAAA 4280  
QY 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGGTGTGTTAGCCAGAGTGAAGAGTTGCTTC 2221  
Db 4279 AGGAGCTCCAATCAGTCTAAGACAAAAAATGGTTGGCTCCCTTACATATGCGACGACA 4220  
QY 2222 TCCGTGATAATGTAATTTGCCGCGATACCCAAAGGACAGACATTCACACCTTTACATTTAGC 2281  
Db 4219 GGTGGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCTGATGGC 4175  
QY 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGATATTTGTTACAACACAGAGCTGATGTGA 2341  
Db 4174 TGCTCATTTGGGACATGTCGCTGCTGCTAACTTTTCTGATGATGATGCTGACCCGAA 4115  
QY 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATATAATGACAGCATCTTACGGGATGTAGA 2401  
Db 4114 TGCTCGAGCTCTCAATGGCTTCACACGCGTGCATATCGCTTGCACAAAAAATGCGATTAA 4055  
QY 2402 TGTAGCAGCTCTACTAATAAAGTATATGATGTGTCATGCGCACGACAAATGGGCTTT 2461  
Db 4054 AATTGTCGAAGTGTCTACTGAAATACCAGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995  
QY 2462 CACACCTTTTGCACCAAGCAGCCCAAGGAGGACAAACAGCTTTGTGCTTTGTTGCTAGC 2521  
Db 3994 CTCACCGCTGCATGCTCGTCTTTTATGGGTGCTATAAACAATGTCATCTATTACTACA 3935  
QY 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581

Db 3934 ACAAGTGCTAATGCAGATGTGGCTACAGTACCGGTGAAGCCCTCTTCAATTAGCTGC 3875  
 QY 2582 A 2582  
 Db 3874 A 3874

## RESULT 9

US-08-847-429A-35

; Sequence 35, Application US/08847429A

; Patent No. 5827692

; GENERAL INFORMATION:

; APPLICANT: Tang, Liang

; APPLICANT: Biehm, E. Scot

; TITLE OF INVENTION: DIOPHILARIA AND BRUGIA ANKYRIN

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: US\$ THEREOF

; NUMBER OF SEQUENCES: 85

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkington Verser, Ph.D.

; ADDRESSEE: Heskia Corporation

; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80525

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: WordPerfect for Windows, Version 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/847,429A

; FILING DATE: 24-APR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Verser, Carol Talkington

; REGISTRATION NUMBER: 37,459

; REFERENCE/DOCKET NUMBER: HW-5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 970/493-7272

; TELEFAX: 970/484-9505

; INFORMATION FOR SEQ ID NO: 35:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5235 nucleotides

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cdna

US-08-847-429A-35

Query Match 2.4%; Score 93; DB 1; Length 5235;

Best Local Similarity 45.0%; Pred. No. 6,6e-14;

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

QY 1392 GGAGCAACATCAATGAAGAGCTAAAGAACTTCTTCACTCTCTGACGTGGCATCTGAG 1451

Db 178 GGCACCGATATCAACACATGCAATGCGAATGCGCTTAATGCAATTCATCTGGCCCTCCAAA 237

QY 1452 AAAGCTCAATAGTGTGTTGAAGTACGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511

Db 238 GAAGTCAATGAAGTGGTCCGCGAAGCTCTGAAAAGAAAGAGAGATGTTGATGCTGCC 297

QY 1512 GATAATCTTGGTCAGACTCTCTACACAGAGTGCATATTTGGTGCATCTACAAACCTGC 1571

Db 298 ACTAGAAAGGTTAACACAGCGTTACATATAGCATTTGGCAGGACAGAACTAATCTGTC 357

QY 1572 GCGTACTCTGAGCTATGGGTGATCCTTAACATTAATCCCTTCAGGGTTTACTGCT 1631

Db 358 ACAGTACTTGTGAAATGGTGTCTAATGTTAAGCTTACATCACTAACCGGTTTACACCA 417

QY 1632 TTACAGATGGGAAATGAAATGTACAGCAACTCTCCAGAGGGTATCTCATAGGTAAT 1691

## RESULT 10

US-08-847-429A-36/c

; Sequence 36, Application US/08847429A

; Patent No. 5827692

; GENERAL INFORMATION:





```

; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-065-474-35

```

Query Match 2.4%; Score 93; DB 3; Length 5235;

Best Local Similarity 45.0%; Pred. No. 6.6e-14;

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

```

QY 1392 GGAGCAACATCAATGAAAGACTAAAGAAATCTTGACTCCTCTGACGCTGGCATCTGAG 1451
D 178 GGCACCCATATCAACATGCAATGCAATGGCTTAAATGCAATGCAATGGCTGCTGCTGCTG 237
QY 1452 AAAGCTCATATGATGTTGTTGAAGTAGTGTGAAGCAATCAAGCAAGAGTTAAATGCTCTG 1511
D 238 GAAGGTCAATCATGAGTGGTCCGGAACTCTGAAAAGAAAGCAGATGTTGATGTGCC 297
QY 1512 GATAATCTTGGTCAGACTTCTACACAGAGCTGCATATTTGTGCTCATCTACAACCTGC 1571
D 298 ACTAGAAAGGTAACACAGCGTTACATATAGCATATTGGCAGGACAAAGAACTAAATGCTC 357
QY 1572 CGCTACTCTGAGCTATGGTGTGATCTTCAATATATATCCCTTACAGGCTTTACTGCT 1631
D 358 ACAGTACTCTTGAATGCTGTAATGTTAAAGTAACTCAATCACTAAACGGTTTACACCA 417
QY 1632 TTACAGATGGGAAATGAAATGTACAGCAACTCTCTCAAGAGGTTATCTCATTAGGTAAT 1691
D 418 CTTTACATGGCTGCACAGAAATCAGAACTCTGTTGACGCTATCTCTTGCCCAAT 477
QY 1692 TCAGAGCAGACAGACAATTTGCTGGAGCTGCAAGGCTGGAGATGT------CGAAAC 1744
D 478 GCCAATCAAGCTTTAAGTACAGAAAGACGGTTTTACGCCACTGGCAGTGTGCTTGAACAA 537
QY 1745 TGTAAAAAACTGTACTGTTTCAGAGTGTCACTGCAGAGACATTTGAAGGGCTCAGTC 1804
D 538 GGTACAGATGCTGTGGTGGCTGTTTGTGTTGAAATGACACGCGCGGAAAGTG-CGCTT 596
QY 1805 TACACCACTTCAATTTTGCAGCTGGGTATACAGAGTGTCTCGTGGTGGAAATATCTGTACA 1864
D 597 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656
QY 1865 GCATGAGCTGATGTCATGCTCAAGATTAAGGAGCGCTTGTACCTTTGACAAATGCATG 1924
D 657 AAATGAGCATAACTCGGATGTGACTTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716
QY 1925 TTCTTATGGCATTATGAAGTTTCAGAACTCTTGTGTTAAACATGGGAGCAGTACTTAAATGT 1984
D 717 TCNCTATGAAATGCAAGCTAGCAACTGCTACTCTGAAAAGGGAGGCAATGTGAATTA 776
QY 1985 AGCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAAAGGAAATATGAAAT 2044
D 777 CCAAGCGAGACATAAACAATAGTCCGTTACAGCTTTCGCAAAATGGGTCGTACAAACAT 836

```

```

QY 2045 TTGCAAACTTCTGCTCCAGCATGTTGCAGACCCCTACCACAAAAAACAGGATGGAATAC 2104
D 837 GGTTCGTTATGTTGGCTCATGGGCGCGTAATTGACTGTGCGCACACGTGATTTACTAAC 896
QY 2105 TCCTTTGGATCTTGTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
D 897 ACCATTACACTGTCTCTCTGCTTCAAGTCATGATCAAGTTGTTGCTTGTGCTTTGCTTGA 956
QY 2162 TGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTTC 2221
D 957 AGGAGCTCCAATCAGTGTCTAAGACAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016
QY 2222 TCCTGATAATCTAAATTCGCCCGGATACCCAGGAGCAGACATTCACACCTTTACATTTAGC 2281
D 1017 GGTGATGATGTTACTGTTGACTATCTC-----ACTCCTCTTCATGTGCG 1061
QY 2282 AGCTGGTTATAATAATTTAGAAAGTTGCAGAGTATTTGTTTACACACGAGCTGATGTCAA 2341
D 1062 TGCTCAATGGGACATGTCGCTGTCGCTAACTTTTGTGCTGATGTAATGCTGACCCGAA 1121
QY 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAATGCAGCATCTTACGGGCGATGTAGA 2401
D 1122 TGCTCGAGCTCTCAATGGCTTCACACGCTGCATATCGCTTGCAAAAAAATCGCATTA 1181
QY 2402 TGTAGCAGCTCTACTATAAAGTATAATGATGATGTTCAATGCCAGGACAAATGGGCTTT 2461
D 1182 AATTGTCGAACTGCTACTGAAATACCAGCTGCAATCGAAGCAACTACTGTAATCCGGTCT 1241
QY 2462 CACACCTTTGCAGAGCAGCGCCAAAAGGAGGACGACAGCTTTGTGCTTTGTTGCTAGC 2521
D 1242 CTCACCGCTGCATGCTGCTGCTTTTATGGTGTCTATAAACATTTCTCATCTATTTACTACA 1301
QY 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACCTTTAGATTTAGTTTC 2581
D 1302 ACAAGTGTCTAATGTCAGATGTTGGCTACAGTACGCGGTGAACGCGCTCTCTATTAGCTGC 1361
QY 2582 A 2582
D 1362 A 1362

```

# RESULT 12

```

US-09-065-474-36/c
; Sequence 36, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459

```

1. *Journal of the American Medical Association*, 1997; 277: 1039-1043.





Qy	1512	GATTAATCTTCGTGACAGCTTCTCTCTACACAGAGTGCATATTTGTGGTCACTATACAAACCTGC	1571
Db	348	ACTAGAAGGTTAACACACGGTTTACATATAGCATCATTTGGCAGGACAGAACTAATCGTC	407
Qy	1572	CGCTACTCTCGAGCTATGGGTGTGATCCTAACATTTATATCCCTTCAGGCGTTTACTGCT	1631
Db	408	ACAGTACTTGTTCAAAATGGTGTATGTTTACGTACAAATCATTAAACGGTTTTTACACCA	467
Qy	1632	TTACAGATGGAAATGAAATGTACAGCAACTCTCCCAAGAGGGTATCTCTATTAGTGAAT	1691
Db	468	CTTTTACATGGCTGCACAAGAAAATCAGGAATCTGTGTGTACGCTATCTTCTGTGCCACAAT	527
Qy	1692	TCAGAGCGACACAGACAATTCCTGGNAGCTGCAAGGCTGGAGATGT-----CGAAAC	1744
Db	528	GCCNATCAAGCTTTTAACTCAGAAGACGGTTTTTACGCCACTTGGCAGTTGCTTTGCAACAA	587
Qy	1745	TGTAAAAAACTGTGTACTGTTTCAGAGTGTCAACTGTCAGAGACATTTGAAGGGGTCAGTC	1804
Db	588	GGTCACGATCGTGTGTCGCTGTTTGTGTAATATGACACGCCGCGGAAGTG-CGCTT	646
Qy	1805	TACACACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAATATCTGTACA	1864
Db	647	GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA	706
Qy	1865	GCATGGAGCTGTGTCATGCTAAAGATAAAGGAGGCCCTGTACCTTGTGCACAATGCATG	1924
Db	707	AAATGACCATTAATTCGGATGTGACCTCGNAAAGCGCTTTACTCCGCTTCATATCGCCGC	766
Qy	1925	TTCTTATGGACATTTAAGTTTGCAGAACTTCTTTTAAACATGGAGCAGTAGTTAATGT	1984
Db	767	TCATATTGAAATGACAAGCTAGCAACTGCTACTCGAAAGGAGCCAAATGTGAATTA	826
Qy	1985	AGCTGATTTATGAAATTTACACCTTTACATGAAGCAGCACAAAGGAAAATATGAAAT	2044
Db	827	CCAAGCGAGACATAACATAAGTCCGTTACCGTTGCACAAAAATGGGGTCGTACAACAT	886
Qy	2045	TTGCAAACTTCTGCTCCAGCATGGTGCAGACCTTACCAAAAAAACAGGATGGAATAC	2104
Db	887	GGTTTCGTTATTCTTGGCTCATGGGCGCTAATTGACTGTGCGACACGTGATTTACTTAAC	946
Qy	2105	TCCTTTGGATCTTGTAAAGATGGAGATACAGAT---ATTCAAGATCTGTTAGGGGAGA	2161
Db	947	ACCATTACATGTGCTCTCTGTTTCAGGTCATGATCAAGTTGTGATTTGTGCTTGAATA	1006
Qy	2162	TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTC	2221
Db	1007	AGGAGCTCCAATCAGTGCTAAGACAAAAATGGTTTGCTCCCTTACATATGCGACACA	1066
Qy	2222	TCCTGATATGTAAATTTGCCGGCATACCAAGGACAGACATTCACACCTTTACATTTAGC	2281
Db	1067	GGTGGATGATGTTACTGTGTTGCACTATCTC-----ACTCCTCTTCATGTGC	1111
Qy	2282	AGCTGGTTATATAATTTAGAAGTTGCAGAGTATTTTGTTCACACAGGACTGATGTGA	2341
Db	1112	TGCTCATTCGGACATGTCCTGTGCTTAACATTTTGTGTGATCGTAATGCTGACCCGAA	1171
Qy	2342	TGCCAAGACAAAGGAGGACTTTTTCCTTTTACATAATGCAGCACTTTTACGGGCGATAGA	2401
Db	1172	TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATGTCGTTGCAAAAAAATCGCATTA	1231
Qy	2402	TGTAGAGCTCTACTAATAAGATATAATGCATGTCTCAATGCCACGACAAATGGGCTTT	2461
Db	1232	AATTGTCGAATCTACTTGAATATCCGCGTGCATCGAAGCAACTACTGAATCCGGTCT	1291
Qy	2462	CACACCTTTGCAGACGACGCCAAAAGGAGACAGACTTTTGTGCTTTGTGTGCTAGC	2521
Db	1292	CTCACCGCTGCATGTGCTGCTTTTATGGTGCTATAAAACATTGTCATCTATTACTACA	1351
Qy	2522	CCATGAGCTGACCCGACTCTTAAAAATCAGAAGGACAAACACCTTTAGATTTAGTTTC	2581
Db	1352	ACAAGTGCTTAATGACAGATGTGGCTTACAGTACGGCGTGAACCGCTCTCATTTAGCTGC	1411
Qy	2582	A	2582





Db 181 CGAAATACAGATGGAGGACAGCATTTGGATTAGCAGATCCATCTGCCAAAGCAGTGCTT 240  
QY 741 ACTGCTGAATATAAGAAAGATGAACCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAAA 800  
Db 241 ACTGGTGAATATAAGAAAGATGAACCTCTAGAAAGTGCAGGAGTGGCAATGAAGAAAA 300  
QY 801 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCAGCAAGTGATGGCAGAAAGTCA 860  
Db 301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCAGCAAGTGATGGCAGAAAGTCA 360  
QY 861 ACTCCATTACATTTGGCAGCAGGATATACAGAGTAAGATTGTACAGCTGTACTGCAA 920  
Db 361 ACTCCATTACATTTGGCAGCAGGATATACAGAGTAAGATTGTACAGCTGTACTGCAA 420  
QY 921 CATGGAGCTGATGTCATCAAGATAAAGTGTATGATCGTACCATTAACAAATGCCCTGT 980  
Db 421 CATGGAGCTGATGTCATCAAGATAAAGTGTATGATCGTACCATTAACAAATGCCCTGT 480  
QY 981 TCTTATGTCATTATGAAGTAAGTAACTGTTGGTCAAGCATGGTGCCTGTGTAAGTCA 1040  
Db 481 TCTTATGTCATTATGAAGTAAGTAACTGTTGGTCAAGCATGGTGCCTGTGTAAGTCA 540  
QY 1041 ATGGACTGTGGCAATTCACCTCTCTCATGAGGCGAGCTTCTAAGAACAGGCTTGAAGTA 1100  
Db 541 ATGGACTGTGGCAATTCACCTCTCTCATGAGGCGAGCTTCTAAGAACAGGCTTGAAGTA 600  
QY 1101 TGTCTCTCTCTTAAGTATGGTGCAGACCCAACTGCTCAATTGTGCACAAATAAAGT 1160  
Db 601 TGTCTCTCTCTTAAGTATGGTGCAGACCCAACTGCTCAATTGTGAAGATAAAGT 660  
QY 1161 GGTATAGACTTGGCTCCACACCCAGTGTAAAGAAAGATTAGCATATGAATTTAAAGGC 1220  
Db 661 GGTATAGACTTGGCTCCACACCCAGTGTAAAGAAAGATTAGCATATGAATTTAAAGGC 720  
QY 1221 CACTGCTGCTCAAGCTGCAGGAGAGTGTACTCGAATCAAAAACATCTCTCT 1280  
Db 721 CACTGCTGCTCAAGCTGCAGGAGAGTGTACTCGAATCAAAAACATCTCTCT 780  
QY 1281 CTGGAATGGTGAATTTCAAGCATCCTCAACACATGAACAGCATGTCATTTGCTGCT 1340  
Db 781 CTGGAATGGTGAATTTCAAGCATCCTCAACACATGAACAGCATGTCATTTGCTGCT 840  
QY 1341 GCATCTCCATATCCAAAGAAAGCAATATGTGAACCTGTGCTAAGAAAGGAGCAAAAC 1400  
Db 841 GCATCTCCATATCCAAAGAAAGCAATATGTGAACCTGTGCTAAGAAAGGAGCAAAAC 900  
QY 1401 ATCAATGAAAGACTAAAGAAATCTTGACTCCTGACGTGGCATCTGAGAAAGCTCAT 1460  
Db 901 ATCAATGAAAGACTAAAGAAATCTTGACTCCTGACGTGGCATCTGAGAAAGCTCAT 960  
QY 1461 AATGATGTTGTTGAAGTGTGTTGAACATGAAGCAAGGTTAATGCTCTGATGAATCTT 1520  
Db 961 AATGATGTTGTTGAAGTGTGTTGAACATGAAGCAAGGTTAATGCTCTGATGAATCTT 1020  
QY 1521 GGTACAGCTTCTACAGAGCTGCATATTTGCTCATCAAAACCTGCGCGCTACTC 1580  
Db 1021 GGTACAGCTTCTACAGAGCTGCATATTTGCTCATCAAAACCTGCGCGCTACTC 1080  
QY 1581 CTGAGCTATGGGTGATGCTCAACATATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1640  
Db 1081 CTGAGCTATGGGTGATGCTCAACATATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140  
QY 1641 GGAATGAAATGTACAGCACTCTCCAAAGGATATCTAATAGGTAATTCAGAGGCA 1700  
Db 1141 GGAATGAAATGTACAGCACTCTCCAAAGGATATCTAATAGGTAATTCAGAGGCA 1200  
QY 1701 GACAGCAATTCCTGGAAGCTCAAGGCTGAGATGTCGAACCTGTAAGAAACCTGTGT 1760  
Db 1201 GACAGCAATTCCTGGAAGCTCAAGGCTGAGATGTCGAACCTGTAAGAAACCTGTGT 1260  
QY 1761 ACTGTTACAGTGTCAACTGCAAGACATTTGAAGGCGCTCAGTCTACACCCTTCATTTT 1820  
Db 1261 ACTGTTACAGTGTCAACTGCAAGACATTTGAAGGCGCTCAGTCTACACCCTTCATTTT 1320

QY 1821 GCAGCTGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTC 1880  
Db 1321 GCAGCTGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGAGCTGATGTC 1380  
QY 1881 CATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCAATGCAATGATGTTCTTATGGACATTA 1940  
Db 1381 CATGCTAAAGATAAAGAGGCGCTTGTACCTTTGCAATGCAATGATGTTCTTACGGACATTA 1440  
QY 1941 GAAAGTGGAGAACTCTCTTAAACATGAGCAGTAGTAAATGATGCTGATTTATGGAAA 2000  
Db 1441 GAAAGTGGAGAACTCTCTTAAACATGAGCAGTAGTAAATGATGCTGATTTATGGAAA 1500  
QY 2001 TTTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTCTCTGCTC 2060  
Db 1501 TTTTACACCTTTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTCTCTGCTC 1560  
QY 2061 CAGCATGGTGCAGACCCCTACCAAAAACACAGGATGGAATACTCTTTGGATCTGTTT 2120  
Db 1561 CAGCATGGTGCAGACCCCTACCAAAAACACAGGATGGAATACTCTTTGGATCTGTTT 1620  
QY 2121 AAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 2180  
Db 1621 AAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCT 1680  
QY 2181 GCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTCTCTCTGATAAATGTAATTTA 2240  
Db 1681 GCCAAGAAAGGTTGTTTAGCCAGAGTGAAGAAAGTTGCTCTCTCTGATAAATGTAATTTA 1740  
QY 2241 CCGCATACCAAGGAGCAGATTTCAACCTTTACATTTAGCAGCTGGTTATATTAATTTA 2300  
Db 1741 CCGCATACCAAGGAGCAGATTTCAACCTTTACATTTAGCAGCTGGTTATATTAATTTA 1800  
QY 2301 GAAAGTGGAGATATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGCAAGAGGAGCA 2360  
Db 1801 GAAAGTGGAGATATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGCAAGAGGAGCA 1860  
QY 2361 CTTATTTCTTTACATATGACAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420  
Db 1861 CTTATTTCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920  
QY 2421 AGTATAATGCATGTCTCAATGCCAGGACAAATGGGCTTTCCACACCTTTGACCAAGCA 2480  
Db 1921 AGTATAATGCATGTCTCAATGCCAGGACAAATGGGCTTTCCACACCTTTGACCAAGCA 1980  
QY 2481 GCCCAAAAGGGAGCAACAGAGCTTTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2540  
Db 1981 GCCCAAAAGGGAGCAACAGAGCTTTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040  
QY 2541 CTTTAAATTCAGGAAGGACAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCT 2600  
Db 2041 CTTTAAATTCAGGAAGGACAAACACCTTTTAGATTTAGTTTTCAGCGGATGATGTCAGCGCT 2100  
QY 2601 CTTCTTGACAGCAGCCATGCCCATCTGCTCTGCTCTGCTTTTACAAGCCTCAAGTGTCTC 2660  
Db 2101 CTTCTTGACAGCAGCCATGCCCATCTGCTCTGCTCTGCTTTTACAAGCCTCAAGTGTCTC 2160  
QY 2661 AATGTTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCATCTAGCCCA 2720  
Db 2161 AATGTTGTGAGAAGCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCATCTAGCCCA 2220  
QY 2721 TCAAGCCTTTCTGACAGCCAGAGCTTTGACAACTTTATCTGGAGGTTTTTCAGAACTGTCT 2780  
Db 2221 TCAAGCCTTTCTGACAGCCAGAGCTTTGACAACTTTATCTGGAGGTTTTTCAGAACTGTCT 2280  
QY 2781 TCAGTAGTTAGTTCAAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAAGAGGTTTCCA 2840  
Db 2281 TCAGTAGTTAGTTCAAGTGAACAGAGGCTGCTCCAGTTTGGAGAAAAGAGGTTTCCA 2340  
QY 2841 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTTGGACTTTGAGCACCTTAATGGAT 2900  
Db 2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTTGGACTTTGAGCACCTTAATGGAT 2400



Db 927 CCAAAACATTCGGAACACTGATGGAAATCAGCCCTGGACCTGGCAGATCCTTTACGAAAA 986  
Qy 732 GCAGTCCCTTACTGCTCAATATTAAGAAAGATGAACCTTTAGAAGTCCAGAGAGTGGCAAT 791  
Db 987 GCTGCTCCTTACAGGTGAATACAAAGAAAGACAGACCTTAGAAGCTGCTAGGAGTGGTAAT 1046  
Qy 792 GAAGAAAAATGATGGCTCTACTACACCATTAATAGTCAACTGACCCACCAAGTATGGC 851  
Db 1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAATGTGAATGGCATGCCATGATGGG 1106  
Qy 852 AGAAGTCAACGCCATTAATTTGGCAGCAGGATATAACAGAGTAAGATTTGACAGCTG 911  
Db 1107 CGAAAGTCACTCCTTTACATCTAGCAGGGGCTACACAGAGTTGCGAATAGTTCACTT 1166  
Qy 912 TTACTGCAACATGGAGCTCATCTCAATGCTAAGATAAAGGTGATCTGCTACCATTAACAC 971  
Db 1167 CTTCTTCAGCATGGTGTGCTGATGTTATGCAAAAGACAAAGGTGGAGCTTGCTCTTCAT 1226  
Qy 972 AATGCTGTTCTTATGCTCATTAAGTAACTGAACCTTTTGGTCAAGCATGGTGCCTG 1031  
Db 1227 AATGCATGTTCAATGACATTAAGTCAAGTCAAGAACTGCTACTAAGCATGGAGCTTGT 1286  
Qy 1032 GTAATGCAATGGACTTGTGGCAATTCATCTCTTCATGAGCAGCTTCTTAAGAACAG 1091  
Db 1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCAAGAGGCTGCTTCCAAAGACCGT 1346  
Qy 1092 GTTGAAGTATGTTCTCTTCTTAACTTATGTTGACAGCCCAACACTGCTCAATTGTCAC 1151  
Db 1347 GTAGAAGTCTGCTTGTGTTACTTAGCCATGGCTGATCTTACGTTAGTCAACTGCCAT 1406  
Qy 1152 AATAAAGTGTATAGACTTGGTCCACACCACAGTTTAAAGAAAGATTAAGCATATGAA 1211  
Db 1407 GGCAAAAGTGTGTGATATGGCTCCAACTCCGGAGCTTAGGAGAGATTGACTTTATGAA 1466  
Qy 1212 TTAAAGGCCACTGTTGCTGCAAGCTGCACGAGAGCTGATGTTACTCGAATCAAAAA 1271  
Db 1467 TTTAAAGGTCATCTTTACTACAGCAGCCAGAGAGAGCTTAGCTAAAGTTAAAAA 1526  
Qy 1272 CATCTCTCTGGAATGTGTAATTTCAAGCATCTTCAACACATGAAACAGCATTTGCAT 1331  
Db 1527 ACATCGCTCTGGAATCATTAATTTCAAAACAGCCAGCTCTCATGAAACAGCACTGCAC 1586  
Qy 1332 TGTGCTGCTGATCTCCATATCCAAAAGAAAGCAAAATATGTAAGTGTGCTAAGAAAA 1391  
Db 1587 TGTGCTGTGGCTCTCTGATCCCAACATCAACAAAGTACAGAAATTTGTTACTTAGAAA 1646  
Qy 1392 GGAGCAACATCAATGAAAGACTAAGAAATTTCTTCACTCTCTGACAGTGGCATCTGAG 1451  
Db 1647 GGAGCAAAATGTTAATGAAAAAATAAAGATTTTCACTACCTCCCTGATGTCAGCCGAA 1706  
Qy 1452 AAAGCTCAATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511  
Db 1707 AGAGCCCAATATGATGTCATGGAAGTTCTGCTAAGCATGGCGCAAGATGAATGACATG 1766  
Qy 1512 GATAATCTTGGTCAGACTTCTCTACACAGAGTGCATATTTGGTGCATCTAACAAGCTGC 1571  
Db 1767 GACACCTTTGGTCAGACTCTTTGATAGAGCGCCCTAGCAGGCCACCTGCAAGACCTGC 1826  
Qy 1572 CGCTACTCTGAGCTPATGGGTGATCTTAACATTAATCTCCTTCAGGCTTTACTGCT 1631  
Db 1827 CGCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCTTACAGGCTTCACAGCA 1886  
Qy 1632 TTACAGATGGGAATGAAATGTACAGCAACTCTTCCAAAGAGGTTATCTCATTAGGTAAT 1691  
Db 1887 GCACAGATGGCAATGAACAGTGCAGCAGATTTCTGAGTGAGAGTACACCTATACGTACT 1946  
Qy 1692 TCAGAGGCAGACAGCAATTTGCTGGAGCTGCAAGGCTGGAGATGTCGAAACTGTAAAA 1751  
Db 1947 TCTGATGTTGATATCGACTCTTTAGAGGATCTTAAAGCTGGAGCTTTGGAAACTGTGAAG 2006  
Qy 1752 AAACCTGTGACTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGGCTGAGTCTACACCA 1811  
Db 2007 CAACTTTGCACTCTCAAAATGTGAATTTAGAGACTTTAGAGGCGCGGCATTTCCACGCC 2066

Qy 1812 CTTCAITTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGA 1871  
Db 2067 TTACACTTCGACAGAGGCTTACACCGCTGTCTGTGTAGAGTACCTGCTACACACGGT 2126  
Qy 1872 GCTGATGTCATGCTTAAAGATAAAGAGAGCCCTTGTACCTTTTGCACAAATGCATGTTCTTAT 1931  
Db 2127 GCCGATGCTCATGCTCCAAAGACAAAGGTGCTTGTGGTCCCTTTCATATATGCTGTTTCATAT 2186  
Qy 1932 GCACATTAAGAGTTGACAGAACTTCTTGTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1991  
Db 2187 GGACACTAATGAGGTGGCTGAGCTTTTAGTAAAGCATGGGCTTCTGTCAATGTGGCGGAC 2246  
Qy 1992 TTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAAGGAAAAATGAAATTTTCCAAA 2051  
Db 2247 TTATGGAATTTTACCCCTCTCCATGAAGCAGCAGCTTAAGAAAGATGTAATCTCAAG 2306  
Qy 2052 CTTCTGCTCCAGCATGGTGCAGACCCCTACCAAAAAACAGGATGGAATACTCCTTTG 2111  
Db 2307 CTCCCTTTTAAACATGGAGCAGATCCAACCTAAAAAGACAGAGATGGAATACACCTTTG 2366  
Qy 2112 GATCTGTTAAAGATGGAGATACAGATAATCAAGATCTGCTTAGGGGAGATGCAGCTTTG 2171  
Db 2367 GATTTGGTAAAGAGGAGACACAGATAATTCAGGACTTACTGAAGGGGATGCTGCTTTG 2426  
Qy 2172 CTAGATGCTGCCAAAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTCTCTGTATAAT 2231  
Db 2427 TTGGATGCTGCCAAGAGGGCTCCCTGGCAGAGTGCAGAAGCTGTACCCACAGAAAT 2486  
Qy 2232 GTAATTTGCGGATACCCCAAGSGACATTCACACCTTTTACATTTAGCAGCTGTTAT 2291  
Db 2487 ATCACTGCGAGACACCCAGGGCAGAAATTCACCCCTCTGCACCTGGCAGCAGCTAT 2546  
Qy 2292 AATAATTTAGAAGTTCACAGTATTTGTACACACGAGAGCTGATGTAATGCCAAGAC 2351  
Db 2547 AATACTGGAAGTAGCTGAATATCTTAGAGCATGGAGCTGATGTTAATGCCCAGGAC 2606  
Qy 2352 AAAGGAGGACTTATTCCTTTTACATAATGACAGCTTTTAGGGCATAGATGTAGACGT 2411  
Db 2607 AAGGGTGGTTAATCTCTTCATATGCGGATCTTATGGGCATGTTGACATAGCGCT 2666  
Qy 2412 CTACTAATAAGTATAATGATGTGTAATGCCACGGCAAAATGGCTTTCACACCTTTG 2471  
Db 2667 TTATGATATAATACACACAGTGTGTAATGCAACAGATAAGTGGGCTTTACTCCCTC 2726  
Qy 2472 CACGAAGCAGCCAAAAGGAGCAACACAGCTTTTGTGCTTAGCCCATGGAGCT 2531  
Db 2727 CATGAAGCAGCCAGAAAGAGGAGCGAGCTGTGCGCCCTCCTCTAGGCAATGGTGA 2786  
Qy 2532 GACCCGACTCTTTAAATCAGGAAGCAACACCTTTTAGATTTAGTTTTCAGCGGATGAT 2591  
Db 2787 GACCCCACTGAAGAACCAGGAAGCCAGACGCTCTGATCTGGCAACAGCTGACGAT 2846  
Qy 2592 GTCAGGCTCTTCTGACACAGCCCATGCCCCATCTGCTCTGCCCTTGTGTTACAGCCT 2651  
Db 2847 ATCAGAGCTTTTCTGATAGATGCCATGCCCGAGAGGCTTACCTACCTGTTTAAACCT 2906  
Qy 2652 CAAGTCTCAATGGTGTGAGAGCCCGAGGACCACTGCAGATGCTCTCTTCAGGTCCA 2711  
Db 2907 CAGG-----CTACTGTAGTGAAGTGCCTCTCTGATCTCACACGCA 2945  
Qy 2712 TCTAGCCCATCAAGCCTTTCTGACGAGCAGAGTCTTGAACACTTATCTGGAGTTTTCAT 2771  
Db 2946 TCCACCCCTCTCTGCTCTCGGCTGCCAGCAGCATAGACAACCTCACTGSCCTTTAGCA 3005  
Qy 2772 GAACGTCTTCAGTAGTTAGTTCAAGTGAACAGAGGCTTCCAGTTTGGAGAAAAAG 2831  
Db 3006 GAGTTGGCCTAGGAGGAGCCTCCAATGAGGGGATGGCGCGCGGAAACAGAAAGAAAG 3065  
Qy 2832 -----GAGGTCCAGGAGTAGAATTTTAGCATAACTCAATTCGTAAGAACTTTGGACTT 2885  
Db 3066 GAAGGAGAAGTGTCTGCTTGNACATGAATATCAGCCATTTCTAAAGAGCCTTGGCCTT 3125



QY	792	GAAGAAAAATGATGGCTCTACTACACACCAATAAATCTCAACTGCCAGCAAGTATGCC	851
Db			
Db	1047	GAAGAAAAACTAATGGCTTTACTGACTCTCTAAATGTGNAITTGCCATGCCAAGTGATGGG	1106
QY	852	AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG	911
Db			
Db	1107	CGAAAGTCGACTCCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCCGAATAGTTCAGCTT	1166
QY	912	TTACTGCAACATGGAGCTGATGTCCATGTCTAAAGATAAAGGTGCATCTGGTACCATTACAC	971
Db			
Db	1167	CTTCTTCAGCATGGTGCTGATGTTCTATGCAAAAGACAAGGTGGACTTGTGGCCTCTTCAT	1226
QY	972	AATGGCTGTTCTATGGTCAATATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCTGT	1031
Db			
Db	1227	AATGCATGTTTCATATGGACATTTAAGTACAGAAACTGCTACTAAAGCATGGAGCTTGT	1286
QY	1032	GTAAATGCAATGGAGCTGTGGCAATTCACCTCTCTCATGAGCAGCTTCTTAAGACAGG	1091
Db			
Db	1287	GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCAGAGGCTGCTTTCGAAGAACCGT	1346
QY	1092	GTTGAAGTATGTTCTTCTCTTAAAGTTATGGTGCAGACCCCAACACTGCTCAATTGTCCAC	1151
Db			
Db	1347	GTAGAAGTCTGCTCTTGTGTTACTTAGCCATGCGCTGATCCTACGTTAGTCAACTGCCAT	1406
QY	1152	AATAAAGTGCTATAGACTTGGCTCCACACACACAGTTAAAAGAAGATTAGCATATGAA	1211
Db			
Db	1407	GGCAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA	1466
QY	1212	TTTAAAGGCCACTCGTTGCTCGAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAA	1271
Db			
Db	1467	TTTAAAGTCTATCTTTACTACAAGCAGCCAGAGAGCAGACTTAGCTTAAAGTTAAAAA	1526
QY	1272	CATCTCTCTCGGAATGGTGNAITTCAGCATCTCTCAACACATGAAACAGCATTTGCAT	1331
Db			
Db	1527	ACACTCGCTCTGGAAATCATTAATTTCAACAACCGCAGTCTCATGTAACAGCAGCTGCAC	1586
QY	1332	TGTGCTGCTGCATCTCCATCTCCCAAAAAAGAAAGAAATATGTGAACCTGTTGCTTAAGAAA	1391
Db			
Db	1587	TGTGCTGGGCTCTCTGCATCCCAACAGTAAACAAGTGACAGAATTTGTTACTTAGAAA	1646
QY	1392	GGACAACATCAATGAAAGACTTAAGAAITTTTGACTCTCTCGCAGTGGCATCTGAG	1451
Db			
Db	1647	GGAGCAATGTTAATGAAAAAATAAGAGTTTCATGACTCCCGTGCATGTTGTCAGCCGAA	1706
QY	1452	AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG	1511
Db			
Db	1707	AGAGCCATAATGATGTCATGGAAGTTCTGCATRAGCATGGCGCAAGATGAATGCATG	1766
QY	1512	GATTAATCTTGGTCAGACTTCTTCACAGAGCTGCATATTTGGTFCATCTACAACCTGC	1571
Db			
Db	1767	GACACCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTTAGCAGGCCACCTGCAGACCTGC	1826
QY	1572	CGCTACTCTCTGAGCTATGGGTGTGATCCCTAACATTTATCCCTTCAGGGCTTTTATGCT	1631
Db			
Db	1827	CGCTCTCTGCTGAGTTAGGCTCTGACCCCTCCCATCATCTCTTACAAGGCTTCCACAGA	1886
QY	1632	TTACAGATGGAAATGAAATGTACAGCACT	1663
Db			
Db	1887	GCACAGATGGCAATGAAGAGTGCAGCAGATTTCTGAGTGTGAGTTACGGCTCTGACCCC	1946
QY	1664	-----	1663
Db	1947	TCCATCATCTCCTTACANGGCTTCACAGCAGCAGATGGCAATGAAGCAGTGCAGCAG	2006
QY	1664	-----	1663
Db	2007	ATTCTGAGTGGTCATTGCTAGATAGTATCTACTTCTAGCCCTTAATGGTGATCTTGA	2066
QY	1664	-----	1663
Db	2067	GACGGAGAGATTAGAAGGAAATCTATCCAGCATGTTCTTCACTGCTCAACATGAAGAGTAC	2126

QY	1664	-----	1663
Db	2127	ACCTATACGTACTTCTGATGTTGATTATCGACTCTTTAGAGGCACTCTAAAAGCTGGAGACTT	2186
QY	1664	-----	1663
Db	2187	GGAAACTGTGAACCAACTTTTGAGCTCTCAAAATGTGAATTTGTAGAGACTTTAGAGGGCG	2246
QY	1664	-----	1664
Db	2247	GCATTCCAGCGCCTTACACTTCCGACGAGCTACAAACAGATACACCTATACATACTTCT	2306
QY	1695	GAGGCAGACAGACAAATGCTGGAAGCTGCAAGGCTGGAGATGTCGAAACTGTGAAAAA	1754
Db	2307	GATGTTGATTATCGACTTTAGAGGCATCTAAGCTGGAGACTTGGAACTGTGAAGCAA	2366
QY	1755	CTGTGTACTGTTTCAGAGTGTCAACTGCAGAGACATTTGAAGGCGCTCAGTCTACACCACTT	1814
Db	2367	CTTTGCAGCTCTCAAAATGTGAATTTGTAGAGACTTAGAGGCGCGCATTTCCAGGCCCTTA	2426
QY	1815	CATTTTGCAGCTGGGTATAACAGAGATGTCGTTGGTGGAAATATCTGCTACAGCATGGAGCT	1874
Db	2427	CACTTTCGACGAGGCTTACAACGCGTGTCTGTTGTAGAGTACTCTGCTACACCGGTGCC	2486
QY	1875	GATGTGCATGCTFAAGATAAAGAGGCGCTTGTAACCTTTGCACAAATGCATGTTCTTATGGA	1934
Db	2487	GATGTCCATGCCAAGACAGAGGTGGCTTGGTGCCCTTTCATAATGCGCTTTTCATATGGA	2546
QY	1935	CATTATGAAGTTGCAGAACTTCTTCTTAAACATGGAGCAGTAGTTTAATGTAGCTGATTTA	1994
Db	2547	CACATATGAGGTGGCTTGAGCTTTTAGTAGGCATGGGCTTCTGTCAATGTGGGGACTTA	2606
QY	1995	TGCAAAATTTACACCTTTTACATGAAGCAGCAGCAAGAGAAATATGAATTTGCAAACTT	2054
Db	2607	TGCAAAATTTACCCCTCTCCATGAAGCAGCAGCTAAAGGAAATATGAATCTGCAAGCTC	2666
QY	2055	CTGCTCCAGCATGGTGCAGACCTTACCACAAAAACAGGGATGGAATFACTCCTTTTGAT	2114
Db	2667	CTTTTAAACATGGAGCAGATCCAACTAAAAAGAACAGAGATGGAATATACACCTTTGGAT	2726
QY	2115	CTTGTTAAAGATGGAGATACAGATATTCAGATCTGCTAGGGAGATGCAGCTTTGGCTA	2174
Db	2727	TGTTGTAAGGAGAGACACAGATATTCAGGACTTACTGAAAGGGATGCTGCTTTGTTG	2786
QY	2175	GATGTCGCCAAGAAGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTCTCCTGATAAATGTA	2234
Db	2787	GATGTCGCCAAGAAGGCTGCTGCGAAGAGTGCAGAAGCTCTGTACCCACAGAGAAATC	2846
QY	2235	AATTCGCGCATACCCAGGCAGACATTCACACCTTTACATTTTAGCAGCTGGTTATAAT	2294
Db	2847	AACCTGCAGACACCCAGGGCAGAAATTCACACCCTCTGCACCTGGCAGCAGGCTATAAT	2906
QY	2295	AATTTAGAAGTTGCAGAGATTTTCTTCAACACCGAGCTGATGTGAATGCCCAAGACAAA	2354
Db	2907	AACCTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCAGGACAAG	2966
QY	2355	GGAGACTTATTCCTTTACATAATGCAGCATCTTTACGGGCATGTAGATGTAGCAGCTCTA	2414
Db	2967	GGTGGTTTAATTCCTCTCTATAATTCGCGCATCTTATGGGCATGTTTGACATAGCGGCTTA	3026
QY	2415	CTAATAAGTATAATGCATGTGTCAATGCCACGACAAATGGGCTTTCACACCTTTGCAC	2474
Db	3027	TTGATAAATACACACGTGTGTAAATTGCAACAGATAAGTGGGCGTTTACTCCCCCTCAT	3086
QY	2475	GAGCAGCCCAAAAGGACGAACACAGCTTTGTGCTTTGTTAGTCCCATGGAGCTGAC	2534
Db	3087	GAGCAGCCCAAGAAAGGACGACGCTGTGCGCCCTCTCTAGCGCATGTGTGCAGAC	3146
QY	2535	CCGACTCTTAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTC	2594
Db	3147	CCACCAATGAAGACCCAGGAAGGCCAGACGCCCTCTGGATCTGCCAACAGCTGACGATAC	3206
QY	2595	AGCGCTCTCTGCAGACAGCCATGCCCCCATCTGCTGCGCTTCTGTTTACAAAGCCTCAA	2654



QY 492 GGGGGCTTATTCCTCTCATATGATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551  
Db 747 GGAGGTCTCATCCCGCTTCATAATGCTGTCTTTTGGCCATGCTGAGGTGTGAGTCTG 806  
QY 552 CTTTTCGGACATGGTCGACAGCCCAATGCTCGAGATAATTGGAAATTATCTCTCTCCAT 611  
Db 807 TTATTTGCGCAAGGAGCTGATCCAAATGCGAGGATAACTGGAACATATACACCTCTGCAT 866  
QY 612 GAAGCTGCAATTAAGGAAAGTTGATGTTTGCATTTGCTGTTACAGATGGAGCTGAG 671  
Db 867 GAAGCTGCTATTAAAGGAAAGATCGATGTTGCTGCTGCGAGCAGCGAGCTGAC 926  
QY 672 CCAACCATCCGAATACAGATGGAAGACAGACATTTGGATTTAGCAGATCCATCTGCCAAA 731  
Db 927 CCAACATTCGGAACACTGATGGAAATCAGCCCTGACCTGSCAGATCCTTCAGCAAAA 986  
QY 732 CGAGTCTCTACTGCTGAATATAAGAAGATGAACCTCTTGAAGAGTCCAGGAGTGCAAT 791  
Db 987 GCTGTCCTTACAGGTGAATACAGAAGACGAACTCCTAGAGCTCTAGGAGTGGTAAT 1046  
QY 792 GAAGAAAAATGATGGCTCTACTCACACCATTAATATGTTCAACTGCCACGCAAGTGATGC 851  
Db 1047 GAAGAAAACTAATGCTTTACTGACTCCTCTAAATGTGAATGGCCATGCAAGTGATGG 1106  
QY 852 AGAAGTCAACTCCATTACATTTGGCAGGAGATATAACAGAGTAAAGATTSTACAGCTG 911  
Db 1107 CGAAGTCTGACTCCTTTACATCTAGCAGCGGCTACACAGAGTTTCCAATAGTTTCACTT 1166  
QY 912 TTTACTGCAATGAGCTGATGCCATGCTAAAGATTAAGGTCATGCTGTACCATTACAC 971  
Db 1167 CTTCTTCAGATGGTCTGATGTTTCATGCAAAAGACAAAGGTGGACTTGTGCCCTCTTCAT 1226  
QY 972 AATGCCCTGTTCTTATGGTCATTATGAAGTAACTGAATTTTGGTCAAGCATGGTGCTGT 1031  
Db 1227 AATGCATGTTTCATATGACATTAATGAATCAGACAGACTGCTACTAAGCATGGAGCTGT 1286  
QY 1032 GTAATATGCAATGGACTGTGGCAATTCATCCTCTTCATGAGCAGCTCTCTAAGAACAG 1091  
Db 1287 GTTAATGCCATGATCTCTGGCAGTTTACTCCACTGACGAGGCTGCTCCRAAGACCGT 1346  
QY 1092 GTTGAAGTATGTTCTCTCTTAACTTTATGGTCAGACCCCAACACTGCTCAATTCGTAC 1151  
Db 1347 GTAGAAGTCTGCTCTTTGTTACTTACCATGGCGCTGATCCCTACGTTAGTCAACTGCCAT 1406  
QY 1152 AATAAAGTCTATAGACTTGGCTCCACACCACAGTTTAAAGAAAGATTAGCATATGAA 1211  
Db 1407 GGCAAAAGTCTGTGGATATGGCTCCAACCTCCGGAGCTTAGGAGAGATTGACTTATGAA 1466  
QY 1212 TTTAAGGCCACTGTTGCTGCAAGCTGACGAGAGCTGATGTTACTCGAATCAAAAAA 1271  
Db 1467 TTTAAAGGTCATCTTTACTACAAGCAGCGCAGAGAAGCAGACTTAGCTAAAGTTAAAAA 1526  
QY 1272 CATCTCTCTGGAAATGTTGAATTTCAAGCATPCTTCAACACATGAAACAGCATGTCAT 1331  
Db 1527 ACNCTCGCTCTGGAATCAATTAATTTCAACACACCCGAGTCTCATGAAACAGCAGCTGAC 1586  
QY 1332 TGTGCTGCTGCATCTCCATATCCAAAAGAAAGCAATATGTGAACGTGTGTTAAGAAAA 1391  
Db 1587 TGTGCTGTGGCCTCTCTGATCCCAACCGTAAACAAGTCAGACGAATTTGTTACTTTAGAAA 1646  
QY 1392 GGAGCAACATCAATGAAAGACTAAGAAATCTTTCACCTCTCTGACAGTGGCATCTGAG 1451  
Db 1647 GGAGCAAAATGTTAAATGAAAAAATAAAGATTTTCATGACCTCCCTGCTGTTCGACCGAA 1706  
QY 1452 AAAGCTCAATATGATGTTGTTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTG 1511  
Db 1707 AGAGCCCATTAATGATCATGGAAGTTCTGCATAAGCATGGGCCAAGATGAATGCACATG 1766  
QY 1512 GATAATCTTTGGTCAGACTTCTCTACACAGAGCTGCAATATTTGGTCACTCAAAACCTGC 1571  
Db 1767 GACACCTTTGGTCAGACTGCTTTTGCATAGAGCGCGCCTAGCAGGCCACCTCGACAGCTGC 1826  
QY 1572 CGCCTACTCTGAGCTATGGGTGTGATCCTAACATTAATCCCTTCAGGGCTTTACTGCT 1631

Db 1827 CGCCTCTCTGTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886  
QY 1632 TTACACATGGGAATGAAATGTACAGCAACTCTCCAAGAGGTTATCTCATTTAGTAAAT 1691  
Db 1887 GCACAGTGGCAATGAAGCAGTGCAGCAGATTTCTGAGTGAGAGTACACATATACCTACT 1946  
QY 1692 TCAGAGGCAGACAGCAATTTGCTGGAGCTGCAAGGCTGGAGATGTGCAAACTGTAAAA 1751  
Db 1947 TCTGATGTTGATTTATCGACTCTTTAGAGGCATCTTAAAGCTGGAGACTTGGAAACTGTGAAG 2006  
QY 1752 AAACCTGTGACTGTTTCAGAGTGTCAACTGCAGACATTTGAAGGGCTCAGTCTACACCA 1811  
Db 2007 CAACTTTGCAGCTCTCAAAATGTGAATTTGTAGACTTTAGAGGCCGGCATTTCCAGCCCC 2066  
QY 1812 CTTTCATTTTGCAGCTGGGTATTAACAGAGTGTCCGTGGTGAATATCTGCTACAGCATGGA 1871  
Db 2067 TTACACTTTCGACAGGCTTACAACCGCTGTCTGTTGTAGAGTACCTGTACACACCGGT 2126  
QY 1872 GCTGATGTGCATGCTAAAGATAAAGAGGCTTGTACCTTTTGCACAATGCTGCTTCAT 1931  
Db 2127 GCGGATGTCCTGCAAGACAGAGGCTGCTTGGTGGCCCTTCATATGCTGTTCATAT 2186  
QY 1932 GGACATTTAAGTTGCGAGAACTTCTTGTTTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1991  
Db 2187 GGACACTATGAGGTGGCTGAGCTTTTAGTAAGCATGGGCTTCTGTCAATGTGGCGGAC 2246  
QY 1992 TTATGAAATTTACACTTTTACATGAAGCAGCAGCAAGAAAGAAATATGAAATTTGCAAA 2051  
Db 2247 TTAAGGAAATTTACCCCTCTCCATGAAGCAGCAGCTTAAAGGAAAGTATGAAATCTCAAG 2306  
QY 2052 CTTCTCTCCAGCATGGTGCAGACCTTACCAAAAAACAGGATGGAATACTCCTTTTG 2111  
Db 2307 CTCCTTTTAAACATGGAGCAGATCCAACCTTAAAGAACAGAGATGGAATACACCTTTG 2366  
QY 2112 GATCTTTTAAAGATGGAGATACAGATATTTCAAGATCTGCTTAGGGAGAGTGCAGCTTTG 2171  
Db 2367 GATTTGGTAAAGGAGGAGACAGATATTCAGAGCTTACTGAAAGGGGATGCTGCTTTG 2426  
QY 2172 CTAGATGCTGCCAAGAGGTTGTTTAGCCAGAGTGAAGAAGTTGTTCTCTCTCTGATAAT 2231  
Db 2427 TTGGATGCTGCCAAGAGGCTGCCCTGGCAAGAGTGCAGAAAGCTCTGACCCACAGAAAT 2486  
QY 2232 GTAAATTTGCGGATACCCCAAGCAGACATTTCAACACCTTTTACATTTAGCAGCTGTTAT 2291  
Db 2487 ATCAACTGCAGAGACACCCAGGCGAGAAATTTCAACCCCTCTGCACCTGGCAGCAGGCTAT 2546  
QY 2292 AATAATTTAGAAGTTGCAGAGTATTTGTTTACAACACGGAGCTGATGTGAATGCCAAGAC 2351  
Db 2547 AATAACTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGAC 2606  
QY 2352 AAAGGAGGACTTATTCCTTTTACATTAATGCAGCATCTTACGG----- 2392  
Db 2607 AAGGTGGTTTAAATTCCTCTTTCATAATGGCGCATCTTTATGGGGGCTGCTTGGCAAGAGTG 2666  
QY 2393 ----- 2392  
Db 2667 CAGAAGCTCTGTACCCACAGAGATAATCAACTGCAGAGACACCCAGGCGAGAAATTCAAACC 2726  
QY 2393 ----- 2392  
Db 2727 CTTCTCCACTGGCAGCAGGCTATAATAACCTGGAAGTAGTGAATATCTTCTAGAGCAT 2786  
QY 2393 ----- 2392  
Db 2787 GGAGCTGATGTTAATGCCCAGGACAGGGTGGTTTAAATTCCTCTTTCATATATGCGGCATCT 2846  
QY 2393 ----- 2392  
Db 2847 TATGGGTAGTAAAAAGTTGGATTCCAAGACCTCTTTCCAGACTTGTGTTAATGATTAATG 2906  
QY 2393 -----GCATGT 2398  
|||||



Db 2907 AGACCATGCATGTGGAATTCGATTAACATTAAGTAAGGCATTATATAAAATGCAAGCATGT 2966  
QY 2399 AGATGTACAGCTCTACTAATAAAGTATAATGCGATGTGTCATGCCAGGACAAATGGC 2458  
Db 2967 TGACATAGCGGCTTTATGATAAAATACACACCGTGTGTAATGCAACAGATAAGTGGC 3026  
QY 2459 TTTCACACTTTGACGAAGCAGCCAAAGGAGGACGAACACAGCTTTGTGCTTTGTGCT 2518  
Db 3027 GTTTACTCCCTCCATGAAGCAGCCGAGAAAGGAAGGACGAGCTGTGCGCCCTCCTCT 3086  
QY 2519 AGCCCATGGAGTACCGAGCTCTTTAAATTCAGGAAGGACAAACACTTTTAGATTTAGT 2578  
Db 3087 AGCGCATGGTCAGAGCCCGACCATGAAGAACCCAGGAAGCCAGAGCCCTCTGGATCTGGC 3146  
QY 2579 TTCAGCGGATGATGTCAGGCTCTCTCT----- 2605  
Db 3147 AACAGCTGACGATATCAGAGCTTTGCGATGTTGACATAGCGGCTTTATGTATAAATACAA 3206  
QY 2606 ----- 2605  
Db 3207 CACGTGTGAATGCAACAGATAAGTGGGCGTTTACTCCCTCCATGAAGCAGCCAGAA 3266  
QY 2606 -----GACAGCAGCATGCCCGCATCTGCTCTGCCCTCTTTTACAAGCCTCAAG----- 2655  
Db 3267 AGGAAGGAGCGAGCTGTGCGCCCTCTCTAGCGCATGGTCAGAGCCCGCCACCATGAAGAA 3326  
QY 2656 -----TGCTCAA 2662  
Db 3327 CCAGGAAGCCAGAGCCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGTAT 3386  
QY 2663 TGGTGTGGAAGCCCGAGGAGCCACTGCG----- 2689  
Db 3387 AGATGCCATGCCCCCAGAGGCGCTTACCTACCTGTTTAAACCTCAGGCTACTGTAGTGAG 3446  
QY 2690 AGATGCTCTCTCTTCAGGTCCATCTAGGCCATCAAGCCCTTTCTGCGAGCCAGCAGCTTGA 2749  
Db 3447 TGCCCTCTCTGATCTACCCAGCATCACCCCTCTGCTCGCTCGGCTCGCAGCAGCATAGA 3506  
QY 2750 CAACCTATCTGGAGTTTTCAGAACTCTCTCAGTAGTTAGTTCAAGTGAACAGAGGG 2809  
Db 3507 CAACCTACTGCGCTTTAGCAGAGTTGGCGTAGGAGGCGCTCCAATGAGGGGATGG 3566  
QY 2810 TGCTTCCAGTTTGGAGAAAGAGG-----GTTCCAGGAGTAGATTTTAGCATAACTCA 2863  
Db 3567 CGCCGCGGAACAGAAAGGAAGGAGAGAGTTGCTGCTTGACATCAATATCAGCCA 3626  
QY 2864 ATTCGTAAGGAATCTTGGACTTTGACCCCTAATGCAATATTTAGAGAGAACAGATCAC 2923  
Db 3627 ATTTCTAAAGCCCTTGGCGCTTGAACACCTTCGGGATATCTTTGAACAGAACAGATTAC 3686  
QY 2924 TTTGATGTATTAGTTGAGATGGGCGACAAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2983  
Db 3687 ACTAGATGTTGGCTGATAGGGTCATGAAGAGTTGAAGAAATAGGCATCAATGCAATA 3746  
QY 2984 TGGACATAGGCAACAACTAATTTAAAGGAGTCGAGAGCTTATCTCCGACACAAAGGTCT 3043  
Db 3747 TGGGCACCCCAACAAATTAATCAAGGAGCTAGAAAGACTCTTTAGCTGGACAAACAGGCAC 3806  
QY 3044 TAACCATATTTAACTTTGAACACCTCTGGTAGTGGACAAATCTTTATAGATCTGCTCC 3103  
Db 3807 CAATCCTTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGTGGAATTTGCTCC 3866  
QY 3104 TGATGATAAGAGTTTCACTGTCTGGAGGAAGAGATGCAAGTACAGTTTCGAGAGCACAG 3163  
Db 3867 AGAGATAAAGAAATATCACTCAAGTGAAGAGATGCAAGTACTATTTTCGAGAACACAG 3926  
QY 3164 AGATGAGGTCTATGCAAGTGGAACTTTCAACAGATACAAATTTCTCAAGATTCAGAAGGT 3223  
Db 3927 AGATGCTGTTATGCTGGCGGATCTTCAACAGATACAAATGTCATTTGAAATTCAAAAAGT 3986  
QY 3224 TTGTACAGAAACTATGGGAAGATACACTCACCGGAGAAAGAGTTTCTGAAGAAAA 3283  
Db 3987 TGTCAACAAGAGTTTGAGGGGCGGTTCTGCCACCGACAGAGAGGAGTGTCTGTAGAGAA 4046

QY 3284 CCACAACCATGCCAATGCAACGAATGCTATTTCATGGTCTCCTTTTGTGAATCAATATAT 3343  
Db 4047 TCACAACCATCACATGACGCGCATGTTGTTTCATGGTTCCTTCCTTCAATTAATGCCATAT 4106  
QY 3344 CCACAAGGCTTTGATGAAAGGCATCGCTACATAGGTGATGTTGGAGCTGGCATTTA 3403  
Db 4107 TCATAAGGCTTTGATGAGCGCATGCATACATAGGAGGAATCTTTGGGCGCGGATTTA 4166  
QY 3404 TTTTGTGAAAACTCTCCAAAAGCAATCAATATGATATGTAATGGAGAGGTACTGG 3463  
Db 4167 TTTTGTGAAAACTCTCCAAAAGCAATCAATATGTTATGTAATGGAGAGGAACAGG 4226  
QY 3464 GTGTCCAGTTTCACAAAGACAGATCTTGTACATTTGCCAGCGCAGCTGCTCTTTGGCGG 3523  
Db 4227 CTGCCCTACACACAGCAGAGTCATGCTATATATGTCACAGACAAATGCTCTCTGTAG 4286  
QY 3524 GGTAACTTTGGGAAGTCTTTCTGCGAGTTCAAGTGAATGAAATGGCAGATTTCTCTCC 3583  
Db 4287 AGTGACCTTTGGGAAATCTCTTCTGCACTTTAGCACCATGAAATGGCCCGCGCTCC 4346  
QY 3584 AGTCACTCACTCACTCACTGCTAGGCCAGTGAATGGCTAGCATTTAGCTCAATATGT 3643  
Db 4347 AGGCACTCACTCACTCACTGCTAGCGGCGCAATGGGCTGGCATATGCTGAATATGT 4406  
QY 3644 TATTTACAGAGGAGACAGGCTTATCTCTGAGTATTTAAATTTACTTACACAGATTTAGAGCC 3703  
Db 4407 CATCTACAGAGGAGACAGCATACCAGAGATATCTTATCACTTACCAGATCATGAAGC 4466  
QY 3704 TGAAG 3708  
Db 4467 AGAAG 4471

## RESULT 5

US-09-964-899-40  
; Sequence 40, Application US/09964899  
; Patent No. US20020174446A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Dalia et al.  
; TITLE OF INVENTION: Identification of Genes Involved in  
; FILE REFERENCE: 4-31612 A  
; CURRENT APPLICATION NUMBER: US/09/964,899  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/236,893  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/298,309  
; PRIOR FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 2409  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-964-899-40

Query Match 31.3%; Score 1192.6; DB 9; Length 2409;

Best Local Similarity 67.4%; Pred. No. 9.9e-314; Indels 1128; Gaps 4;  
Matches 2372; Conservative 0; Mismatches 19;

QY 225 ATGTGCGGTCCGCTGCGCGCGGGGAGCGGCTTGGCGAGCGCGCGCGCGCGCGGCCC 284  
Db 1 ATGTGCGGTCCGCTGCGCGCGGGGAGCGGCTTGGCGAGCGCGCGCGCGCGCGGCCC 60  
QY 285 GTGAGCGCGCGCGCGCGAGAGCTGTCGAGCGTGGCGCAACGGGGAGCTGGGAACGAGTC 344  
Db 61 GTGAGCGCGCGCGCGCGAGAGCTGTCGAGCGTGGCGCAACGGGGAGCTGGGAACGAGTC 120  
QY 345 AAGAGCTGGTGACGCTGGAAGGTGAACAGCGCGACACGCGCGCGCGCGCGGAAATCCACC 404  
Db 121 AAGAGCTGGTGACGCTGGAAGGTGAACAGCGCGCGCGCGCGCGGAAATCCACC 180

QY 405 CGCTGCACTTCGCCCGCAGGTTTGGCGGGAAGACGTAGTTGAATATTGCTTTCAGAAAT 464  
Db 181 CGCGTCACCTTCGCCCGCAGGTTTGGCGGGAAGACGTAGTTGAATATTGCTTTCAGAAAT 240  
QY 465 GGTGCAAAATGTCACACGACGTGATGATGGGCGCTTATTCCCTTCATCAATGCATGCTCT 524  
Db 241 GGTGCAAAATGTCACACGACGTGATGATGGGCGCTTATTCCCTTCATCAATGCATGCTCT 300  
QY 525 TTTGGTCACTGCAAGTAGTCAATCTCCTTTTGGCAGACATGGTGCACACCCCAATGCTCGA 584  
Db 301 TTTGGTCACTGCAAGTAGTCAATCTCCTTTTGGCAGACATGGTGCACACCCCAATGCTCGA 360  
QY 585 GATAATTGAATATTACTCCTCTCCATGAAGCTGCAATTTAAAGGAAGATGATGTTTGC 644  
Db 361 GATAATTGAATATTACTCCTCTCCATGAAGCTGCAATTTAAAGGAAGATGATGTTTGC 420  
QY 645 ATTGTGCTCTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGAAGGACAGCA 704  
Db 421 ATTGTGCTCTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGAAGGACAGCA 480  
QY 705 TTTGGATTTAGCAGATCCATCTGCCAAGCAGTGTCTTACTGGTGAATAT 752  
Db 481 TTTGGATTTAGCAGATCCATCTGCCAAGCAGTGTCTTACTGGTGAATAT 540  
QY 753 -----AAGAAAGATGAACCTTTAGAAAGTGCAGAGTGGCAATGAAGAAAATGATG 806  
Db 541 TATTCAGGAAGCCCTGTAAGAAACAACTTGCAGGAGTGGCAATGAAGAAAATGATG 600  
QY 807 GCTCTACTCACACCATTTAATGTCACCTGCCACGCAAGTGTGGCAGAAAGTCAACTCCA 866  
Db 601 GCTCTACTCACACCATTTAATGTCACCTGCCACGCAAGTGTGGCAGAA 648  
QY 867 TTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTACTGCAACATGGA 926  
Db 649 ----- 648  
QY 927 GCTGATGTCATGCTAAAGATTAAGGTGATCTGGTACCATTACACAATGACCTGTTCTTAT 986  
Db 649 ----- 648  
QY 987 GGTCAATTATGAATTAACCTGAACTTTTGGTCAAGCATGGTGCTGTGTAATGCAATGGAC 1046  
Db 649 -----AAGCATGGTGCTGTGTAATGCAATGGAC 678  
QY 1047 TTGTGGCAATTCACTCCTCTTCATGAGGAGCTTCTAAGAACAGGGTTGAAGTATGTTCT 1106  
Db 679 TTGTGGCAATTCACTCCTCTTCATGAGGAGCTTCTAAGAACAGGGTTGAAGTATGTTCT 738  
QY 1107 CTTCTCTTAAGTTATGGTGCAGACCCCAACACTGCTCAATGTGTCACAATAAAGTCTCTATA 1166  
Db 739 CTTCTCTTAAGTTATGGTGCAGACCCCAACACTGCTCAATGTGTCACAATAAAGTCTCTATA 798  
QY 1167 GACTTGGCTCCACACCACTGTAAGAAAGATTTAGCATATGAATTTAAAGGCCACTCG 1226  
Db 799 GACTTGGCTCCACACCACTGTAAGAAAGATTTAGCATATGAATTTAAAGGCCACTCG 858  
QY 1227 TTGCTGCAAGCTGCAGAGACCTGATGTTACTCGAATCAAAAACATCTCTCTCTGAAA 1286  
Db 859 TTGCTGCAAGCTGCAGAGACCTGATGTTACTCGAATCAAAAACATCTCTCTCTGAAA 918  
QY 1287 ATGGTGAATTTCAAGCATCTCAACACATGAACAGCATTTGATGCTGCTGCTCATCT 1346  
Db 919 ATGGTGAATTTCAAGCATCTCAACACATGAACAGCATTTG 960  
QY 1347 CCATATCCAAAAGAAAGCAATATGTGAAGTGTGCTTAAGAAAGGACAAACATCAAT 1406  
Db 961 ----- 960  
QY 1407 GAAAAGACTAAGAAATCTTGACTCTCTGCACGTGGCATCTGAGAAAGCTCATTAATGAT 1466  
Db 961 ----- 960  
QY 1467 GTTGTGAAGTAGTGGTGAACATGAAGCAAGGTTAATGCTCTGGATAATCTTGGTCAG 1526

Db 961 ----- 960  
QY 1527 ACTTCTCTACACAGAGCTGCATATTTGGTTCATCTTACAAACCTGCCCTTACTCTCAGC 1586  
Db 961 ----- 960  
QY 1587 TATGGTGTGATCCTTAACATTTATATCCCTTACGGGCTTTACTCTCTTACAGATGGGAAAT 1646  
Db 961 ----- 960  
QY 1647 GAAAATGTACAGCAACTCTCTCCAAGAGGTATCTCATTAGGTAAATTCAGAGGACAGACA 1706  
Db 961 ----- 960  
QY 1707 CAATTGCTGGAAGCTCCAAGGCTGGAGATGTGGAACCTGTAAAAAACTGTGTACTGTT 1766  
Db 961 -----AAACTGTGTACTGTT 975  
QY 1767 CAGAGTGTCAACTGCAGACACATTTGAAGGGCGTCACTCTACACCACTTCATTTTGCAGCT 1826  
Db 976 CAGAGTGTCAACTGCAGACACATTTGAAGGGCGTCACTCTACACCACTTCATTTTGCAGCT 1035  
QY 1827 GGGTATAACAGAGTGTCCGTGTGGAAATATCTCTACAGCATGGAGCTGATGTGCATGCT 1886  
Db 1036 GGGTATAACAGAGTGTCCGTGTGGAAATATCTCTACAGCATGGAGCTGATGTGCATGCT 1095  
QY 1887 AAAGATTAAGGAGGCGCTTGTACTTTGCACAATGCATGTTCTTATGGACATTTATGAAGTT 1946  
Db 1096 AAAGATTAAGGAGGCGCTTGTACTTTGCACAATGCATGTTCTTATGGACATTTATGAAGTT 1155  
QY 1947 GCAGAACTCTTTGTTAAACATGGAGCAGTGTAAATGTAGCTGATTTATGGAAATTTTACA 2006  
Db 1156 GCAGAACTCTTTGTTAAACATGGAGCAGTGTAAATGTAGCTGATTTATGGAAATTTTACA 1215  
QY 2007 CTTTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTCTCTGCTCCAGCAT 2066  
Db 1216 CTTTACATGAAGCAGCAGCAAAAGAAATATGAAATTTGCAAACTCTCTGCTCCAGCAT 1275  
QY 2067 GGTGCAGACCTTACCACCAAAACAGGGATGCAATACCTTGGATCTTGTAAAGAT 2126  
Db 1276 GGTGCAGACCTTACCACCAAAACAGGGATGCAATACCTTGGATCTTGTAAAGAT 1335  
QY 2127 GGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCGCAAG 2186  
Db 1336 GGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGAGCTTTGCTAGATGCTGCGCAAG 1395  
QY 2187 AAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATTAATGTAATTTGCCGCGAT 2246  
Db 1396 AAGGTTGTTTAGCCAGAGTGAAGAAGTTGCTTCTCTGATTAATGTAATTTGCCGCGAT 1455  
QY 2247 ACCCAAGGCAGACATTTCAACACCTTTACATTTAGCAGCTGGTTATAAATTTAGAAGTT 2306  
Db 1456 ACCCAAGGCAGACATTTCAACACCTTTACATTTAGCAGCTGGTTATAAATTTAGAAGTT 1515  
QY 2307 GCAGAGTATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGACAAGAGGAGCTTATT 2366  
Db 1516 GCAGAGTATTTGTTTACAACAGGAGCTGATGTAATGCCCAAGACAAGAGGAGCTTATT 1575  
QY 2367 CCTTTACATTAATGCAGCATCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAGTAT 2426  
Db 1576 CCTTTACATTAATGCAGCATCTTACGGGATGTAGATGTAGCAGCTCTACTAATAAGTAT 1601  
QY 2427 AATGCAATGTCTCAATGCCACGACAAATGGGCTTTACACACCTTTGCACGAGCAGCCCAA 2486  
Db 1602 ----- 1601  
QY 2487 AAGGACGAACACAGCTTTGTGCTTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAAA 2546  
Db 1602 ----- 1601  
QY 2547 AATCAGGAAGCAACACCTTTAGATTTAGTTTTCAGCGGATGATGTACGCGCTCTTCTG 2606

Db 1602 ----- 1601  
Qy 2607 ACAGCAGCATGCCCCCATCTGCTCTGCCCTCTTTGTTACAAAGCCTCAAGTGCTCAATGGT 2666  
Db 1602 ----- 1601  
Qy 2667 GTGAGAGCCCGAGGACCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGC 2726  
Db 1602 ----- 1601  
Qy 2727 CTTTCTGCGCCAGCAGTCTTGACAACCTTATCTGGAGTCTTTTCAGAACTGTCTTCAGTA 2786  
Db 1602 ----- 1601  
Qy 2787 GTTAGTCAAGTGGACAGAGGGTCTTCCAGTTTGGAGAAAAGGAGGTTCCAGGAGTA 2846  
Db 1602 ----- 1601  
Qy 2847 GATTTTAGCATAACTCAATTCTGAAGAACTCTTGAGCTTGAGCACCTTAATGGATATATT 2906  
Db 1602 ----- 1601  
Qy 2907 GAGAGAAACAGATCAGCTTTGATGTATTAAGTGGGCGCACAAAGAGCTGAAGGAG 2966  
Db 1602 -----GATCAGCTTGGATGATTAGTTGAGATGGGCGCACAAAGAGCTGAAGGAG 1650  
Qy 2967 ATTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATC 3026  
Db 1651 APTGGAATCAATGCTTATGGACATAGGCACAACTAATTAAGGAGTCGAGAGACTTATC 1710  
Qy 3027 TCCGCACAAAGGCTTAAACCCATATTAACCTTTGAACACCTCTGTTAGTGGAAACAATT 3086  
Db 1711 TCCGCACAAAGGCTTAAACCCATATTAACCTTTGAACACCTCTGTTAGTGGAAACAATT 1770  
Qy 3087 CTTATAGATCTCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGT 3146  
Db 1771 CTTATAGATCTCTCTCTGATGATAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAGT 1830  
Qy 3147 ACAGTTCGAGACACAGAGATGGAGTCTATGAGTGGAACTTCAACAGATACAAATATT 3206  
Db 1831 ACAGTTCGAGACACAGAGATGGAGTCTATGAGTGGAACTTCAACAGATACAAATATT 1890  
Qy 3207 CTCAGATTCAGAAGGTTTGAACAGAACTATGGGAAGATACACTCACCGGAGAAAA 3266  
Db 1891 CTCAGATTCAGAAGGTTTGAACAGAACTATGGGAAGATACACTCACCGGAGAAAA 1950  
Qy 3267 GAAGTTTCTGAAGAAACACACAACTATGCCAATGAACGAATGCTATTTTCATGGGTCTCCT 3326  
Db 1951 GAAGTTTCTGAAGAAACACACAACTATGCCAATGAACGAATGCTATTTTCATGGGTCTCCT 2010  
Qy 3327 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAGGCATGCGCTACATAGTGGTATG 3386  
Db 2011 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAGGCATGCGCTACATAGTGGTATG 2070  
Qy 3387 TTTGGAGCTGGCATTTATTTTCTGCTGAAAACTCTTCCAAAGCAATCAATATGATATGGA 3446  
Db 2071 TTTGGAGCTGGCATTTATTTTCTGCTGAAAACTCTTCCAAAGCAATCAATATGATATGGA 2130  
Qy 3447 ATTGGAGGAGTACTGGGTGTCAGTTTCACAAAGCAGATCTGTACATTTGCCACAGG 3506  
Db 2131 ATTGGAGGAGTACTGGGTGTCAGTTTCACAAAGCAGATCTGTACATTTGCCACAGG 2190  
Qy 3507 CAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTTCCCTGCAGTTCAGTGCATGAAA 3566  
Db 2191 CAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTTCCCTGCAGTTCAGTGCATGAAA 2250  
Qy 3567 ATGGCAGATTTCTCTCCAGGTCATCACTCAGTCACTGTGTAGGCCAGTGAATGGCCTA 3626  
Db 2251 ATGGCAGATTTCTCTCCAGGTCATCACTCAGTCACTGTGTAGGCCAGTGAATGGCCTA 2310  
Qy 3627 GCATTAGCTGAATATGTTTATTTACAGAGGAGACAGGCTTATCCTCAGTATTTAATTACT 3686  
Db 2311 GCATTAGCTGAATATGTTTATTTACAGAGGAGACAGGCTTATCCTCAGTATTTAATTACT 2370

Qy 3687 TACCAGATTATGAGGCTCAAGGTATGTCGATGGATAA 3725  
Db 2371 TACCAGATTATGAGGCTCAAGGTATGTCGATGGATAA 2409

RESULT 6

US-09-833-381-841  
; Sequence 841, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: ROBISON, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 841  
; LENGTH: 1069  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(1069)  
; OTHER INFORMATION: n = A,T,C or G  
; US-09-833-381-841

Query Match 12.3%; Score 467.6; DB 10; Length 1069;  
Best Local Similarity 98.8%; Pred. No. 8,1e-117;  
Matches 492; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 1562 CTCTCCAAAGGGTATCTCATTTAGTAAATTTCAGAGCAGACAGACAATTTGCTGGAAGCT 1721  
Db 567 CCCTCTAGAGGTATCTCTATTAGGTAATTCAGAGCAGACAGACAATTTGCTGGAAGCT 626  
Qy 1722 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1781  
Db 627 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 686  
Qy 1782 AGAGACATTTGAAGGGCTGAGTCTACACCACTTCATTTTCAGCTGGGTATACAGAGTG 1841  
Db 687 AGAGACATTTGAAGGGCTGAGTCTACACCACTTCATTTTCAGCTGGGTATACAGAGTG 746  
Qy 1842 TCCCTGTGTGAATATCTGCTACAGCATGGAGCTGATGCTGCTAAAGATAAAGGAGGC 1901  
Db 747 TCCGTGTGTGAATATCTGCTACAGCATGGAGCTGATGCTGCTAAAGATAAAGGAGGC 806  
Qy 1902 CTTGTACCTTTGCAATGCTGCTTATGGACATTTGAAAGTTGAGAACTTCTTTGTT 1961  
Db 807 CTTGTACCTTTGCAATGCTGCTTATGGACATTTGAAAGTTGAGAACTTCTTTGTT 866  
Qy 1962 AAACATGAGCAGTAGTATTAATGAGTACCTTACACCTTTACATGAAGCA 2021  
Db 867 AAACATGAGCAGTAGTATTAATGAGTACCTTACACCTTTACATGAAGCA 926  
Qy 2022 GCAGCAAAAGCAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACC 2081  
Db 927 GCAGCAAAAGCAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCCTACA 986  
Qy 2082 AAAAAAACA -GGGATGGAATACTCCTTTTGGATCTTTGTTTAAAGATGAGATACAGATAT 2140  
Db 987 AAAAAAACAAGGGATGGAATACTCCTTTTGGATCTTTGTTTAAAGATGAGAT -CAGATAT 1045  
Qy 2141 TCAAGATCTGCTAGGGG 2158  
Db 1046 TCAAGATCTGCTAGGGG 1063

RESULT 7

US-09-833-381-1153

; Sequence 1153, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090a1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1153  
; LENGTH: 465  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1153

Query Match 12.1%; Score 461; DB 10; Length 465;  
Best Local Similarity 100.0%; Pred. No. 2.8e-115;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 572 CCCCAATGCTCGAGATAATGGAAATTAATCTCTCCATGAAGCTGCAATTAAGGAAA 631  
Db 5 CCCCAATGCTCGAGATAATGGAAATTAATCTCTCCATGAAGCTGCAATTAAGGAAA 64  
QY 632 GATTGATGTTTGGCATTGCTGTTACAGCATGGAGCTGAGCCAAACCATCGGAAATACAGA 691  
Db 65 GATTGATGTTTGGCATTGCTGTTACAGCATGGAGCTGAGCCAAACCATCGGAAATACAGA 124  
QY 692 TGGAGGACAGCATGTTGATTTACAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 751  
Db 125 TGGAGGACAGCATGTTGATTTACAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 184  
QY 752 TAAGAAAGATGAATCTCTAGAAATGCGAGAGTGCGCAATGAAGAAATGATGGCTCT 811  
Db 185 TAAGAAAGATGAATCTCTAGAAATGCGAGAGTGCGCAATGAAGAAATGATGGCTCT 244  
QY 812 ACTCACACCATTAATGTCCTCAACTGCCAGCAAGTGATGGCAGAAAGTCAACTCCATTACA 871  
Db 245 ACTCACACCATTAATGTCCTCAACTGCCAGCAAGTGATGGCAGAAAGTCAACTCCATTACA 304  
QY 872 TTTGGCAGCAGGATATAACAGATGAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 931  
Db 305 TTTGGCAGCAGGATATAACAGATGAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 364  
QY 932 TGTCATGCTAAAGATAAAGTGATCTGTTACCATTAACCAATGCTGTTCTTATGGTCA 991  
Db 365 TGTCATGCTAAAGATAAAGTGATCTGTTACCATTAACCAATGCTGTTCTTATGGTCA 424  
QY 992 TTATGAAGTAAGTGAATCTTTGGTCAAGCATGGTGCTGTG 1032  
Db 425 TTATGAAGTAAGTGAATCTTTGGTCAAGCATGGTGCTGTG 465

## RESULT 8

US-10-040-739-1179  
; Sequence 1179, Application US/10040739  
; Patent No. US20020173635A1  
; GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth  
McCoy, John  
Lavallie, Edward  
Racie, Lisa  
Merberg, David  
Treach, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS  
NUMBER OF SEQUENCES: 1519  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge

STATE: Massachusetts  
COUNTRY: U.S.A  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/040,739  
FILING DATE: 07-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/036,520  
FILING DATE: 03-JUN-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 1179:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1179:  
US-10-040-739-1179

Query Match 8.5%; Score 323; DB 9; Length 353;  
Best Local Similarity 99.7%; Pred. No. 7.7e-78;  
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3252 ACTCACCGGAGAAAGAGTTTCTGAAGAAAACACCAACCATGCCAATGAACGAATGCTA 3311  
Db 21 ACTCACCGGAGAAAGAGTTTCTGAAGAAAACACCAACCATGCCAATGAACGAATGCTA 80  
QY 3312 TTTTCATGGGTCTCTCTTTTGGAATGCAATATCCAAAGGCTTTGATGAAGGATCGG 3371  
Db 81 TTTTCATGGGTCTCTCTTTTGGAATGCAATATCCAAAGGCTTTGATGAAGGATCGG 140  
QY 3372 TACATAGGTGATGTTGAGCTGGCATTATTTGCTGAAAACCTCTCCAAAAGCAAT 3431  
Db 141 TACATAGGTGATGTTGAGCTGGCATTATTTGCTGAAAACCTCTCCAAAAGCAAT 200  
QY 3432 CAATATGATATGGAATTTGAGGAGGTACTGGGTGCCAGTTCACAAAGACAGATCTTGT 3491  
Db 201 CAATATGATATGGAATTTGAGGAGGTACTGGGTGCCAGTTCACAAAGACAGATCTTGT 260  
QY 3492 TACATTTGCCACAGCGAGCTGCTCTTTTGGCGGTAACTTGGGAAGTCTTTTCCCTGCAG 3551  
Db 261 TACATTTGCCACAGCGAGCTGCTCTTTTGGCGGTAACTTGGGAAGTCTTTTCCCTGCAG 320  
QY 3552 TTCAGTGCATGAAAATGGCAGAT 3575  
Db 321 TTCAGTGCATGAAAATGGCAGAT 344

## RESULT 9

US-09-841-835-12  
; Sequence 12, Application US/09841835  
; Patent No. US20020076795A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor









APPLICANT: Raju, Jeyaseelan  
TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
FILE REFERENCE: MNI-068CP2  
CURRENT APPLICATION NUMBER: US/09/947,199  
PRIOR FILING DATE: 2001-09-05  
PRIOR APPLICATION NUMBER: 60/111,938  
PRIOR FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: 09/291,839  
PRIOR FILING DATE: 1999-04-14  
PRIOR APPLICATION NUMBER: 09/458,457  
PRIOR FILING DATE: 1999-12-10  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2505  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(2505)  
US-09-947-199-3

Query Match 1.9%; Score 74.2; DB 10; Length 2505;  
Best Local Similarity 55.1%; Pred. No. 1.4e-09;  
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
QY 1806 ACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGCTGGTGAATATCTGCTACAG 1865  
DB 406 ACTGCCCTCCATATTGCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGTTGCAA 465  
QY 1866 CATGGAGCTGATGCTGCTAAAGATTAACAGAGTGTCCGCTGGTGAATATCTGCTACAG 1925  
DB 466 CATGGAGCTGATGCTGCTAAAGATTAACAGAGTGTCCGCTGGTGAATATCTGCTACAG 525  
QY 1926 TCTTATGGACATTTATGAAGTTCGCAACTTCTTGTAAACATGGAGCAGTAGTAAATGTA 1985  
DB 526 TACTATGGACATGAACAGGTAACCTGCTTCTTTTGAATTTGGTCTGATGTAATGTA 585  
QY 1986 GCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAATATGAAAT 2045  
DB 586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATT 645  
QY 2046 TGCAAACTTCTGCTCCAGCATGG 2068  
DB 646 GCAAACTCTTGATGGAGAGG 668

RESULT 15  
US-09-947-199-1  
Sequence 1, Application US/09947199  
Patent No. US20020127684A1  
GENERAL INFORMATION:  
APPLICANT: Raju, Jeyaseelan  
TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
FILE REFERENCE: MNI-068CP2  
CURRENT APPLICATION NUMBER: US/09/947,199  
PRIOR FILING DATE: 2001-09-05  
PRIOR APPLICATION NUMBER: 60/111,938  
PRIOR FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: 09/291,839  
PRIOR FILING DATE: 1999-04-14  
PRIOR APPLICATION NUMBER: 09/458,457  
PRIOR FILING DATE: 1999-12-10  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3025  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS

LOCATION: (48)..(2552)  
US-09-947-199-1

Query Match 1.9%; Score 74.2; DB 10; Length 3025;  
Best Local Similarity 55.1%; Pred. No. 1.7e-09;  
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
QY 1806 ACACCACTTCATTTTGCAGCTGGGTATACAGAGTGTCCGCTGGTGAATATCTGCTACAG 1865  
DB 453 ACTGCCCTCCATATTGCTACAATAGCTGCCACCTAGAGGCTGCTGATGCTGTTGCAA 512  
QY 1866 CATGGAGCTGATGCTGCTAAAGATTAACAGAGTGTCCGCTGGTGAATATCTGCTACAG 1925  
DB 513 CATGGAGCTAAATGCTCAATATTCAAGATGAGGATTTTTCACCTCCATTTGATGTCAGCG 572  
QY 1926 TCTTATGGACATTTATGAAGTTCGCAACTTCTTGTAAACATGGAGCAGTAGTAAATGTA 1985  
DB 573 TACTATGGACATGAACAGGTAACCTGCTTCTTTTGAATTTGGTCTGATGTAATGTA 632  
QY 1986 GCTGATTTATGGAATTTTACACCTTTTACATGAAGCAGCAGCAAAAGGAAAAATATGAAAT 2045  
DB 633 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATT 692  
QY 2046 TGCAAACTTCTGCTCCAGCATGG 2068  
DB 693 GCAAACTCTTGATGGAGAGG 715

Search completed: February 12, 2003, 03:32:11  
Job time : 215.419 secs



```

FEATURES
  source
    Helix Research Institute.
    Location/Qualifiers
      1..838
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="PLACE2000021"
        /clone_lib="PLACE2"
        /tissue_type="placenta"
        /note="Vector: pME18SFL3"
BASE COUNT      260 a 143 c 203 g 227 t      5 others
ORIGIN
Query Match      20.6%; Score 784.2; DB 9; Length 838;
Best Local Similarity 98.0%; Pred. No. 5e-171;
Matches 823; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
QY 2734 CAGCCAGCAGCTTTGACAACTTATCTGGAGTTTTCAGAACTGCTTCAGTAGTAGTT 2793
Db 1 CAGCCAGCAGCTTTGACAACTTATCTGGAGTTTTCAGAACTGCTTCAGTAGTAGTT 60
QY 2794 CAAGTGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGATTTA 2853
Db 61 CAAGTGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGTTCCAGGAGTAGATTTA 120
QY 2854 GCATAACTCAATTCGTAAAGAACTCTGGAGTTCAGACCTTAATGGATATATTTGAGAGAG 2913
Db 121 GCATAACTCAATTCGTAAAGAACTCTGGAGTTCAGACCTTAATGGATATATTTGAGAGAG 180
QY 2914 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAGGAGCTCAAGGAGATTGAA 2973
Db 181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAGGAGCTCAAGGAGATTGAA 240
QY 2974 TCAATGCTTATGACATAGGCACAACTAATTAAGAGTTCGAGAGACTTATCTCCGGAC 3033
Db 241 TCAATGCTTATGACATAGGCACAACTAATTAAGAGTTCGAGAGACTTATCTCCGGAC 300
QY 3034 AACAAAGTCTTAACCCATATTAACTTTGACACCTCTGGTAGTGGAACTTATCTTATAG 3093
Db 301 AACAAAGTCTTAACCCATATTAACTTTGACACCTCTGGTAGTGGAACTTATCTTATAG 360
QY 3094 ATCTGCTCTCTGATGATTAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTC 3153
Db 361 ATCTGCTCTCTGATGATTAAGAGTTTTCAGTCTGTGGAGGAAGAGATGCAAGTACAGTTC 420
QY 3154 GAGACACAGAGATGAGGTATGACAGTGGATCTTCAACAGATACAATATTCTCAAGA 3213
Db 421 GAGACACAGAGATGAGGTATGACAGTGGATCTTCAACAGATACAATATTCTCAAGA 480
QY 3214 TTCAGAAAGTGTGTAACAGAAACTATGGAAAGATACACTACCCGAGAAAGAGTTC 3273
Db 481 TTCAGAAAGTGTGTAACAGAAACTATGGAAAGATACACTACCCGAGAAAGAGTTC 540
QY 3274 CTGAAGAAACCAACACATGCGCAATGAACGAATGCTATTTCATGGTCTCCTTTTGTGA 3333
Db 541 CTGAAGAAACCAACACATGCGCAATGAACGAATGCTATTTCATGGTCTCCTTTTGTGA 600
QY 3334 ATGCAATTTATCCCAAGGCTTTGATGAAGGATCGGTACATAGTGGTATGTTGGAG 3393
Db 601 ATGCAATTTATCCCAAGGCTTTGATGAAGGATCGGTACATAGTGGTATGTTGGAG 660
QY 3394 CTGGCAATTTATTTGCTGAAAACCTCTTCCAAAAGCAATCAATATGATATGAAATGGAG 3453
Db 661 CTGGCAATTTATTTGCTGAAAACCTCTT -CAAAAGCAATCAATATGATATGAAATGGAN 719
QY 3454 GAGGTACTGGGTGTCAGTTTCACAAAGACAGATCTTGTGTACATTTGCCACAGCAGCTGC 3513
Db 720 GAGT -CTGGGTGTCAGTTTCACAAAGACCGATCTTGGTACATTTGCCACAGCAGCTGC 778
QY 3514 TCTTTTGTG--CCGGTAACTCTGGAAAGTCTTCTCGTGCAGTTTCAGTGCATCAAAATGCG 3571
Db 779 TCTTTTGGCCGGGTAACTCTGGAAAGTCTTCTCGTGCAGTTTCANTGCAATGAAATGCG 838

```

RESULT 2  
 BM457025  
 LOCUS  
 DEFINITION AGENCOURT\_6406851 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5583257  
 5', mRNA sequence.  
 ACCESSION BM457025  
 VERSION BM457025.1 GI:18506065  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1128)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12346 Row: a Column: 18  
 High quality sequence stop: 507.  
 Location/Qualifiers  
 1..1128  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5583257"  
 /clone\_lib="NIH\_MGC\_92"  
 /tissue\_type="embryonal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; oligo-qt primed.  
 Average insert size 2.5 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH-MGC Library."  
 BASE COUNT 328 a 297 c 232 g 271 t  
 ORIGIN  
 Query Match 17.6%; Score 672.2; DB 13; Length 1128;  
 Best Local Similarity 97.7%; Pred. No. 6.1e-145;  
 Matches 724; Conservative 0; Mismatches 13; Indels 4; Gaps 4;  
 QY 1976 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 2035  
 Db 1 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 60  
 QY 2036 ATATGAAATTTGCAAACTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAAACAGGGA 2095  
 Db 61 ATATGAAATTTGCAAACTCTGCTCCAGCATGGTCAGACCCCTACCAAAAAAACAGGGA 120  
 QY 2096 TGAATACTCTCTTTGGATCTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAG 2155  
 Db 121 TGAATACTCTCTTTGGATCTTTAAAGATGGAGATACAGATATTCAGATCTGCTTAG 180  
 QY 2155 GGCAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 2215  
 Db 181 GGCAGATGAGCTTTGCTAGATGCTGCCAAGAGGGTTGTTAGCCAGAGTGAAGAAGTT 240  
 QY 2216 GTCTTCTCTGATATGTAATTTGAAATTTGCCGATACCCAGGACAGATTCACACCTTTTACA 2275  
 Db 241 GTCTTCTCTGATATGTAATTTGAAATTTGCCGATACCCAGGACAGATTCACACCTTTTACA 300  
 QY 2276 TTATGAGCTGGTATATAATTTAGAAAGTTCAGAGTATTTGTTTACAACAGCAGCTGA 2335  
 Db 301 TTATGAGCTGGTATATAATTTAGAAAGTTCAGAGTATTTGTTTACAACAGCAGCTGA 360  
 QY 2336 TGTGAATGCCAAGACAAAGGAGGACTTATTCCTTTACATATATGACAGATCTTACGGGCA 2395  
 Db 361 TGTGAATGCCAAGACAAAGGAGGACTTATTCCTTTACATATATGACAGATCTTACGGGCA 2395



source  
1. .691  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4537138"  
/clone\_lib="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 218 a 122 c 165 g 186 t  
ORIGIN

Query Match 16.9%; Score 643.8; DB 12; Length 691;  
Best Local Similarity 99.3%; Pred. No. 1.8e-138;  
Matches 678; Conservative 0; Mismatches 2; Indels 3; Gaps 3;  
QY 1725 AAGCGTGAGATGCGAACTGTAAAAAAGTGTACTGTTCAGAGTGTCAACTGCAGA 1784  
Db 10 AAGCGTGAGATGCGAACTGTAAAAAAGTGTACTGTTCAGAGTGTCAACTGCAGA 69  
QY 1785 GACATTGAAGGCGTCAGCTACACCCTTCATTGTTCAGCTGGGTATAACAGAGTGTC 1844  
Db 70 GACATTGAAGGCGTCAGCTACACCCTTCATTGTTCAGCTGGGTATAACAGAGTGTC 129  
QY 1845 GTGTGGAATATCTGTACAGCATGGAGCTGATGTGCAATGCTAAAGATAAAGGAGCGCTT 1904  
Db 130 GTGTGGAATATCTGTACAGCATGGAGCTGATGTGCAATGCTAAAGATAAAGGAGCGCTT 189  
QY 1905 GTACCTTTGCACATGCATGTTCTTATGGACATTATGAAGTTGCAGACCTTCTTTGTTAAA 1964  
Db 190 GTACCTTTGCACATGCATGTTCTTATGGACATTATGAAGTTGCAGACCTTCTTTGTTAAA 249  
QY 1965 CATGAGCAGTAGTAAATGATGAGCTGATTATGAGTAAATTTACACCTTTACATGAAGCAGCA 2024  
Db 250 CATGAGCAGTAGTAAATGATGAGCTGATTATGAGTAAATTTACACCTTTACATGAAGCAGCA 309  
QY 2025 GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCA-TGGTGCAGACCCCTACCAA 2083  
Db 310 GCAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCACTGGTGCAGACCCCTACCAA 369  
QY 2084 AAAAAACAGGATGGAATATCTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 2143  
Db 370 AAAAAACAGGATGGAATATCTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 429  
QY 2144 AGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTAGCCAG 2203  
Db 430 AGATCTGCTTAGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAGGTTGTTAGCCAG 489  
QY 2204 AGTGAAGAAGTTGCTTCTCTGATA-ATGTAATTTGCCCGCATACCCAAAGGCAGACATT 2262  
Db 490 AGTGAAGAAGTTGCTTCTCTGATAATGTAATTTGCCCGCATACCCAAAGGCAGACATT 549  
QY 2263 CAACACCTTTACATTTAGCAGCTGGTTATTAATTAATTTAGAGTTGCGAGTATTTGTTAC 2322  
Db 550 CAACACCTTTACATTTAGCAGCTGGTTATTAATTAATTTAGAGTTGCGAGTATTTGTTAC 608  
QY 2323 AACACGAGCTGATGTGAATGCCCAAGACAAAGGAGGACFTATTCTTTACATAATGCAG 2382  
Db 609 AACACGAGCTGATGTGAATGCCCAAGACAAAGGAGGACFTATTCTTTACATAATGCAG 668  
QY 2383 CATCTTTACGGGCATGTAGATGTA 2405  
Db 669 CATCTTTACGGGCATGTAGATGTA 691

RESULT 5  
AL601027 759 bp mRNA linear EST 14-AUG-2001  
LOCUS DKFp313G2239\_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone  
DEFINITION DKFp313G2239 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

BASE COUNT  
ORIGIN

Query Match  
Best Local Similarity  
Matches

Conservative  
Mismatches  
Indels  
Gaps

QY 2393 GCATGTAGATGTAGCAGCTCTACTATAAAGTAAATGCAATGTGTCATGCGACGGACAA 2452  
Db 72 GCATGTANATGTAGCAGCTCTACTATAAAGTAAATGCAATGTGTCATGCGACGGACAA 131  
QY 2453 ATGGGCTTTTACACCTTTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2512  
Db 132 ATGGGCTTTTACACCTTTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 191  
QY 2513 GTTGTAGCCCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGAGGAGGAGGAGGAGG 2572  
Db 192 NTTTTGGCCCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGAGGAGGAGGAGGAGG 251  
QY 2573 TTTAGTTTCAGCGGATGATGTACGCGCTCTTCTGACAGCAGCCTATGCCCATCTGCTCT 2632  
Db 252 TTTAGTTTCAGCGGATGATGTACGCGCTCTTCTGACAGCAGCCTATGCCCATCTGCTCT 311  
QY 2633 GCCTCTTGTGTACAAGCTCAAGTGTCAATGTTGTGAGAGGCCAGGAGGAGGAGGAGGAGG 2692  
Db 312 GCCTCTTGTGTACAAGCTCAAGTGTCAATGTTGTGAGAGGCCAGGAGGAGGAGGAGGAGG 371  
QY 2693 TGTCTCTCTTTCAGGTCCTATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGTCTTGACAA 2752  
Db 372 TGTCTCTCTTTCAGGTCCTATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGTCTTGACAA 431  
QY 2753 CTTATCTGGGAGTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAACAGAGGAGGAGG 2812  
Db 432 CTTATCTGGGAGTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAACAGAGGAGGAGG 491  
QY 2813 TTCACCTTTGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAG 2872  
Db 492 TTCACCTTTGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAG 551

AL601027  
AL601027.1 GI:15164533  
EST.  
human.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 759)  
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann  
S.  
EST (Duesterhoeft, et al.)  
Unpublished (1999)  
Contact: Duesterhoeft A  
MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by Oligen (Hilden/Germany) within the cDNA sequencing  
consortium of the German genome project.  
No sl sequence available.  
This clone (DKFp313G2239) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..759

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFp313G2239"  
/clone\_lib="313 (synonym: hlcc2)"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

BASE COUNT 217 a 156 c 175 g 201 t 10 others

Query Match 16.8%; Score 642.8; DB 9; Length 759;  
Best Local Similarity 97.2%; Pred. No. 3.3e-138;  
Matches 669; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

QY 2393 GCATGTAGATGTAGCAGCTCTACTATAAAGTAAATGCAATGTGTCATGCGACGGACAA 2452  
Db 72 GCATGTANATGTAGCAGCTCTACTATAAAGTAAATGCAATGTGTCATGCGACGGACAA 131  
QY 2453 ATGGGCTTTTACACCTTTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2512  
Db 132 ATGGGCTTTTACACCTTTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 191  
QY 2513 GTTGTAGCCCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGAGGAGGAGGAGGAGG 2572  
Db 192 NTTTTGGCCCATGGAGCTGACCGGACTCTTAAATAATCAGGAAGGAGGAGGAGGAGGAGG 251  
QY 2573 TTTAGTTTCAGCGGATGATGTACGCGCTCTTCTGACAGCAGCCTATGCCCATCTGCTCT 2632  
Db 252 TTTAGTTTCAGCGGATGATGTACGCGCTCTTCTGACAGCAGCCTATGCCCATCTGCTCT 311  
QY 2633 GCCTCTTGTGTACAAGCTCAAGTGTCAATGTTGTGAGAGGCCAGGAGGAGGAGGAGGAGG 2692  
Db 312 GCCTCTTGTGTACAAGCTCAAGTGTCAATGTTGTGAGAGGCCAGGAGGAGGAGGAGGAGG 371  
QY 2693 TGTCTCTCTTTCAGGTCCTATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGTCTTGACAA 2752  
Db 372 TGTCTCTCTTTCAGGTCCTATCTAGCCCATCAAGCCTTTCTGACGCCAGCAGTCTTGACAA 431  
QY 2753 CTTATCTGGGAGTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAACAGAGGAGGAGG 2812  
Db 432 CTTATCTGGGAGTTTTCAGAACTGCTTTCAGTAGTTAGTTCAAGTGAACAGAGGAGGAGG 491  
QY 2813 TTCACCTTTGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAG 2872  
Db 492 TTCACCTTTGGAGAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAG 551

```

QY 2873 GAATCTTGGACTCAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 2932
|||||
Db 552 GAATCTTGGACTCAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 611
|||||
QY 2933 ATTAGTTGAGATGGGACAGGAGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAG 2992
|||||
Db 612 ANNAGTTGAGATGGGACAGGAGCTGAAGGAGATTGGAATCAATGCTTTATGGACATAG 671
|||||
QY 2993 GCACAACTAATTTAAA-GGAGTCGAGAGACTTATCTCCGACACAAAG-GTCTTTAACCCTA 3050
|||||
Db 672 GCACAACTAATTTAAAAGGAGTCGAGAGACTTATCTCCGACACAAAGGGGCTTTAACCCTA 731
|||||
QY 3051 TATTAACTTTGAACACTCTGTGGTAGTG 3078
|||||
Db 732 TATTAACTTTGAACACTCTGTGGNAGAG 759
|||||

RESULT 6
B0885764 950 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8746924 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6392888
5', mRNA sequence.
ACCESSION B0885764
VERSION B0885764.1 GI:22277782
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL 1 (bases 1 to 950)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13883 row: h column: 09
High quality sequence stop: 641.
FEATURES
source
1..950
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6392888"
/clone.lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: otocysts; Vector: pCMV-SPORT6.1.cdb;
Site:1: EcoRV; Site:2: NotI; Cloned unidirectionally.
Primer: Oligo dT. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
BASE COUNT 280 a 195 c 226 g 245 t 4 others
ORIGIN
Query Match 16.6%; Score 634.4; DB 14; Length 950;
Best Local Similarity 86.4%; Pred. No. 3.4e-136;
Matches 759; Conservative 0; Mismatches 111; Indels 8; Gaps 5;

QY 2500 AGCTTTGTGCTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAAATCATCAGGAGGAC 2559
|||||
Db 1 AGCTTTGTGCTTGTGCTAGCCCATGGAGCTGACCCGACTCTTAAATCATCAGGAGGAC 59
|||||
QY 2560 AAACACCTTTAGATTTTACGGGATGATCTCAGCCGCTCTCTGACAGCAGCCATGC 2619
|||||
Db 60 AAACACCTTTAGATTTTACGGGATGATCTCAGCCGCTCTCTGACAGCAGCCATGC 119
|||||
QY 2620 CCCCATCTGCTGCTGCTCTTGTGTACAGCCCTCAAGTGTCTCAATGGTGTGAGAGCCCGAG 2679
|||||

```

```

Db 120 CCCCCCTCTGCTGCTAGCTGCTACAAACCTCAAGTGTGCGGTGAGGGGCCCGG 179
|||||
QY 2680 GAGCCACTCCAGATGCTCTCTCTCAGGTCCATCTAGCCCATCAAGCCCTTTCTGCGAGCA 2739
|||||
Db 180 GAGCCACTCCAGATGCTCTCTCAGGTCCATCTAGCCCATCAAGCCCTTTCTGCGAGCA 239
|||||
QY 2740 GCAGTCTTTCAGCAACTTATCTGGGAGTTTTCAGAACTGTCTTCAGTAGTGTCAAGTG 2799
|||||
Db 240 GCAGCTTCACAACTTATCTGGCAGCTTCTCGAACTGTCCGAGTGTGTAGTTCAGTG 299
|||||
QY 2800 GAACAGAGGCTGCTTCCAGTTTGGGAAAAGAGAGTTCAGGAGTAGATTTTACATAA 2859
|||||
Db 300 CAGCAGAAGGTGCTACTGCTTTTGCAAAAGAAAGAGAGATTTCAGGAATTCGATTTAGTATAA 359
|||||
QY 2860 CTCATTTCTGAAGAACTTGGACTTTCAGCACTTATGATATATTTGAGAGAGACAGA 2919
|||||
Db 360 CTCAGTTTCATAGAACTTGGACTTTCAGCACTTATGATATATTTGAGAGAGACAGA 419
|||||
QY 2920 TCACCTTTGGATGTTATTTAGTTGAGATGGGACAAAGGAGTGAAGGAGTTGGAATCATG 2979
|||||
Db 420 TTACCTTAGATGCTCTAGTTGAAATGGCCACAAAGAACTGAAAGAGATTGGAATCAATG 479
|||||
QY 2980 CTTATGACATAGGCACAACTAATTTAAAGAGTTCAGAGACTTATCTCGGACACAAAG 3039
|||||
Db 480 CCTATGACATAGGCACAACTAATTTAAAGAGTTCAGAGACTTATCTCGGACACAAAG 539
|||||
QY 3040 GTCTTAAACCCATTTTAACTTTTCAACACCTCTGTTAGTGAACAAATTTCTTATAGATCTGT 3099
|||||
Db 540 GTCTTAAACCCATTTTAACTTTTCAACACCTCTGTTAGTGAACAAATTTCTTATAGATCTGT 599
|||||
QY 3100 CTCTCATGATGATGAGAGTCTTTCAGTCTGTGGAGGAGAGATGCAAGTTCAGAGAGC 3159
|||||
Db 600 CTCTCATGATGATGAGAGTCTTTCAGTCTGTGGAGGAGAGATGCAAGTTCAGAGAGC 659
|||||
QY 3160 ACAGAGATGAGGATGATGAGTGTGGAATCTTCAACAGATACATATTTCAAGATTCAGA 3219
|||||
Db 660 ACAGAGATGAGGATGATGAGTGTGGAATCTTCAACAGATACATATTTCAAGATTCAGA 719
|||||
QY 3220 AGGTTTCTTAAACAACTATGGGAAAGATACACTCACCGG-AGAAAAGAGTTTCTGAA 3278
|||||
Db 720 AGGTTTCTTAAACAACTATGGGAAAGATATACACACCGGAGAAAAGAGTTTCTGAA 779
|||||
QY 3279 GAAACCCACACCATGC-CAATGAACGAATCTTATTTTCATGGGTCT---CCTTTTGTGAA 3334
|||||
Db 780 GAAACCCACACCATGC---CCTTTTCATGGGTCTTATTTTCATGGGTCTCCTTTTGTGAA 839
|||||
QY 3335 TGCAATTATCCACAAAGG--CCTTTGATGAAGGATGC 3370
|||||
Db 840 TGCATTATCCATAAGGGCTTTTTCATGAAGGATGC 877
|||||

RESULT 7
B0875327 648 bp mRNA linear EST 30-MAY-2001
LOCUS RCI-BF0623-120200-011-d04 BF0623 Homo sapiens cDNA, mRNA sequence.
DEFINITION B0875327
ACCESSION B0875327
VERSION B0875327.1 GI:14252306
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 648)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

```

MEDLINE  
COMMENT

20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC1-BT0623-120  
200-011-d000423-2000-02-12&tl=1)  
Seq primer: puc 18 forward  
High quality sequence start: 4  
High quality sequence stop: 648.

FEATURES  
source

1..648  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="BT0623"  
/dev\_stage="Adult"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESIES PCR (U.S. Letters Patent application No. 196  
/716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 200 a 118 c 157 g 172 t 1 others

Query Match 15.9%; Score 605; DB 12; Length 648;

Best Local Similarity 98.7%; Pred. No. 1.8e-129;

Matches 620; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1663 TCCTCAAGAGGGTATCTCATTAGTAATTCAGAGCAGACACAAATTCCTGGAGCTG 1722

Db 23 TCTCACAAGGGTATCTCATTAGG--ATTCAGAGCAGACACAAATTCCTGGAGCTG 80

QY 1723 CAAAGCTGGAGATGTCGAACTGTAAAAAACTGTACTGTTCAGAGTGTCAACTGCA 1782

Db 81 CAAAGCTGGAGATGTCGAACTGTAAAAAACTGTACTGTTCAGAGTGTCAACTGCA 140

QY 1783 GAGACATTGAAGGGCTCAGCTACACCACTTCATTTTCAGCTGGGTATACAGAGTGT 1842

Db 141 GAGACATTGAAGGGCTCAGCTACACCACTTCATTTTCAGCTGGGTATACAGAGTGT 200

QY 1843 CCGTGTGGAAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAGATAAAGGAGGCC 1902

Db 201 CCGTGTGGAAATATCTGCTACAGCATGGAGCTGATGTCATGCTAAGATAAAGGAGGCC 260

QY 1903 TTGTACCTTTGCACATGTCATGCTTTATGACATTTATGAAGTTGCAGAACTTCTTGTTA 1962

Db 261 TTGTACCTTTGCACATGTCATGCTTTATGACATTTATGAAGTTGCAGAACTTCTTGTTA 320

QY 1963 AACATGGACAGTATGTTATGTTAGTGTATTTATGAAATTTACACCTTTACATGAAGCAG 2022

Db 321 AACATGGACAGTATGTTATGTTAGTGTATTTATGAAATTTACACCTTTACATGAAGCAG 380

QY 2023 CAGCAAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGCTGCAGACCCCTACCA 2082

Db 381 CAGCAAAAAGGAAATATGAAATTTGCAAACTTCTGCTCCAGCATGCTGCAGACCCCTACCA 440

QY 2083 AAAAAACAGGATGGAATATCTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 2142

Db 441 AAAAAACAGGATGGAATATCTCTTTGGATCTTTGTTAAAGATGGAGATACAGATATTC 500

QY 2143 AAGATCTGCTTAGGGAGATGACGCTTTTGCTAGATGCTGCCAAGAGGCTTTTAGCCA 2202

Db 501 AAGATCTGCTTAGGGAGATGACGCTTTTAGCTAGATGCTGCCAAGAGGCTTTTAGCCA 560

QY 2203 GAGTGAAGAAGTTGTCTTCTCTGTATAATGTAAATTTGCCGCGATACCCAAAGGCAGACATT 2262

## Db

561 GAGTGAAGAAGTTGTCTTCTCTGTATAATGTAATTCGCGATACCCAGGACGACATT 620

## QY

2263 CAACACCTTTACATTTAGCAGCTGGTTA 2290

## Db

621 CAACACCTTTACATATAGCAGCTGGTTA 648

## RESULT 8

BE220005/c

## LOCUS

DEFINITION

BE220005

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

622 bp mRNA linear EST 03-JUL-2000  
hV65H03.x1 NCI-CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3178325 3'  
similar to TR:095271.095271 TRF1-INTERACTING ANKYRIN-RELATED  
ADP-RIBOSE POLYMERASE. ; mRNA sequence.

BE220005  
BE220005.1 GI:8907323  
EST.  
human.

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 622)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-femail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL, send email to:  
info@image.lnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 474.

## FEATURES

## source

1..622  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NCI-CGAP\_Lu24"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a  
modified polylinker; Plasmid DNA from the normalized  
library NCI-CGAP\_Lu5 was prepared, and ss circles were  
made in vitro. Following HAP purification, this DNA was  
used as tracer in a subtractive hybridization reaction.  
The driver was PCR-amplified cDNAs from a pool of 5,000  
clones made from the same library (cloneIDs  
1414920-1417991 and 1520904-1522439). Subtraction by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 179 a 130 c 128 g 185 t

## Query Match

Best Local Similarity 15.5%; Score 593.2; DB 10; Length 622;

Matches 604; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 3183 GGAATCTTCAACAGATACATAATTTCTCAAGATTCAGAGGTTTGTAAACAGAACTATCG 3242

Db 622 GCAATCTTCAACAGATACATAATTTCTCAAGATTCAGAGGTTTGTAAACAGAACTATCG 563

QY 3243 GAAAGATACACTCAGCGGAGAAAGAGTTTCTGAAGAAACCAACCATGCCAATGAA 3302

Db 562 CAAAGATACACTCAGCGGAGAAAGAGTTTCTGAAGAAACCAACCATGCCAATGAA 503

QY 3303 CCAATGCTATTTCATGGGTCTCTCTTTTGTGAATGCAATATCCACAAAGGCTTTGATGAA 3362

Db 502 CGAATGCTATTTCATGGGTCTCTCTTTTGTGAATGCAATATCCACAAAGGCTTTGATGAA 443



QY 3363 AGGCATCGGTACATAGGTGGTATGTTGGAGCTGGCATTTATTTTCTGTAAGAACTCTTCC 3422  
 Db 442 AGGCATCGGTACATAGGTGGTATGTTGGAGCTGGCATTTATTTTCTGTAAGAACTCTTCC 383  
 QY 3423 AAAAGCAATCAATATGATATGGAATGGAGAGGTACTGGGTGTCAGGTTCCAGTTCAACAAGAC 3482  
 Db 382 AAAAGCAATCAATATGATATGGAATGGAGAGGTACTGGGTGTCAGGTTCCACAAGAC 323  
 QY 3483 AGATCTGTTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCT 3542  
 Db 322 AGATCTGTTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCT 263  
 QY 3543 TTCCTGAGTTCAGTGAATGGAATGGCAATCTCTCCAGTGCATCACTCAGTCACT 3602  
 Db 262 TTCCTGAGTTCAGTGAATGGAATGGCAATCTCTCCAGTGCATCACTCAGTCACT 203  
 QY 3603 GGTAGGCCAGTGAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662  
 Db 202 GGTAGGCCAGTGAATGGCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACAG 143  
 QY 3663 GCTATCTCTGAGTATTTAAATTTACTTACAGATTTATGAGGCTGAGGTATGGTCGATGGA 3722  
 Db 142 GCTATCTCTGAGTATTTAAATTTACTTACAGATTTATGAGGCTGAGGTATGGTCGATGGA 83  
 QY 3723 TAAATAGTATTTTAAAGAACTAAATCCACTGAACCTTAAATATCAATCAAAAGCAGTGGC 3782  
 Db 82 TAAATAGTATTTTAAAGAACTAAATCCACTGAACCTTAAATATCAATCAAAAGCAGTGGC 23  
 QY 3783 CTCTACGTTTACTCCTTTGCT 3804  
 Db 22 CTCTACGTTTACTCCTTTGCT 1  
 RESULT 9  
 BE467629/c  
 LOCUS  
 DEFINITION  
 ADP-RIBOSE POLYMERASE. ; mRNA sequence.  
 BE467629  
 BE467629.1 GI:9513404  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 616)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 402.  
 Location/Qualifiers  
 1..616  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3213021"  
 /clone\_lib="NCI-CGAP\_Lu24"  
 /tissue\_type="carcinoid"  
 /lab\_host="DH108"  
 /note="Organ: Lung; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; plasmid DNA from the normalized  
 library NCI-CGAP\_Lu5 was prepared, and ss clones were  
 made in vitro. Following HAP purification, this DNA was  
 used as tracer in a subtractive hybridization reaction.  
 The driver was PCR-amplified cDNAs from a pool of 5,000  
 clones made from the same library (clonoids  
 1414920-1417991 and 1520904-1522439). Subtraction by Bento  
 Soares and M. Fatima Bonaldo. "

BASE COUNT 179 a 130 c 124 g 183 t  
 ORIGIN

Query Match 15.3%; Score 584; DB 10; Length 616;  
 Best Local Similarity 96.8%; Pred. No. 1.3e-124;  
 Matches 596; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 3189 TTCACAGATACAAATATCTCAAGATTTCAGAGGTTTGAACAAAGAACTATGGGAAGA 3248  
 Db 616 TTCACAGATACAAATATCTCAAGATCATATAGTTTGCACCACTATGGATAGA 557  
 QY 3249 TACACTCACCGGAGAAAAGATTTCTGAAGAAAACCAACCATGCCAATGAACGAATG 3308  
 Db 556 TACACTCACCGGAGACAAGAAAGTTTTTGAAGAAAACCCACCATGCCAATGACGAATG 497  
 QY 3309 CTATTTTCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAAGGCTTTGATGAAGGCAT 3368  
 Db 496 CTATTTTCATGGTCTCCTTTTGTGAATGCAATTTATCCAAAAGGCTTTGATGAAGGCAT 437  
 QY 3369 GCGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTCTGAAACCTCTTCCAAAAGC 3428  
 Db 436 GCGTACATAGTGGTATGTTTGGAGCTGGCATTTATTTTGTCTGAAACCTCTTCCAAAAGC 377  
 QY 3429 AATCAATATCATATGCAATTTGGAGAGGTPACTGGGTGTCAGTTTCCAAAAGCAGATCT 3488  
 Db 376 AATCAATATCATATGCAATTTGGAGAGGTPACTGGGTGTCAGTTTCCAAAAGCAGATCT 317  
 QY 3489 TGTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAGTCTTTCCTG 3548  
 Db 316 TGTTACATTTGCCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAGTCTTTCCTG 257  
 QY 3549 CAGTTTCACTGCAATGAAATGGCAATTTCTCTCCAGGTCATCACTCAGTCACTGGTAGG 3608  
 Db 256 CAGTTTCACTGCAATGAAATGGCAATTTCTCTCCAGGTCATCACTCAGTCACTGGTAGG 197  
 QY 3609 CCCAGTGTAAATGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACAGCTTAT 3668  
 Db 196 CCCAGTGTAAATGCCCTAGCATTTAGCTGAATATGTTATTTACAGAGGAGAACAGCTTAT 137  
 QY 3669 CCTCAGTATTTAATTACTTACCAGATTTAGGCTCGAAGGTATGTCGATGGATAATA 3728  
 Db 136 CCTCAGTATTTAATTACTTACCAGATTTAGGCTCGAAGGTATGTCGATGGATAATA 77  
 QY 3729 GTTATTTTAAAGAACTAATTCCTCACTGAACCTTAAATATCAATCAAAAGCAGTGGCTCTAC 3788  
 Db 76 GTTATTTTAAAGAACTAATTCCTCACTGAACCTTAAATATCAATCAAAAGCAGTGGCTCTAC 17  
 QY 3789 GTTTTACTCCTTTGCT 3804  
 Db 16 GTTTTACTCCTTTGCT 1

RESULT 10  
 BE910901  
 LOCUS  
 DEFINITION  
 601661978F1 NCI-CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3962251 5',  
 mRNA sequence.  
 BE910901  
 BE910901.1 GI:10407885  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 956)









Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagareishvili, R., Williams, J., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)

# TITLE JOURNAL COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812  
Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. J. J. Ferrer in vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco  
High quality sequence stop: 429.

## FEATURES source

Location/Qualifiers  
1. .553  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE: 5778810"  
/clone\_lib="Human insulinoma"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/notes="Organ: pancreas; Vector: pBluescript SK-; Site\_1: XhoI; Site\_2: EcoRI; Constructed with lambda ZAPII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol  
(http://genome.wustl.edu/est/Lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

BASE COUNT 165 a 107 c 121 g 160 t  
ORIGIN

Query Match 14.4%; Score 549; DB 14; Length 553;  
Best Local Similarity 100.0%; Pred. No. 1.7e-116;  
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3260 GAGAAAGAGATTCTGAAGAAACACCAACCATGCCAATGAACGAATGCTATTTCATGG 3319  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
5 GAGAAAGAGAGATTCTGAAGAAACACCAACCATGCCAATGAACGAATGCTATTTCATGG 64  
QY 3320 GTCTCCTTTTGTGAATGCATTTATCCACAAAGGCTTTGTATGAAGGCATCGGTACATAGG 3379  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
65 GTCTCCTTTTGTGAATGCATTTATCCACAAAGGCTTTGTATGAAGGCATCGGTACATAGG 124  
QY 3380 TGGTATCTTTGGAGCTGGCATTATTTTGTCTGAAACACTCTTCCAAAGCAATCAATATGT 3439  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
125 TGGTATCTTTGGAGCTGGCATTATTTTGTCTGAAACACTCTTCCAAAGCAATCAATATGT 184  
QY 3440 ATATGGAATTTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTTGTACATTTG 3499  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
185 ATATGGAATTTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTTGTACATTTG 244  
QY 3500 CCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCCTGCGAGTTCAGTGC 3559  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
245 CCACAGCAGCTGCTCTTTTGGCGGGTAACCTTGGGAAAGTCTTTCCTGCGAGTTCAGTGC 304  
QY 3560 AATGAAATGGCACAATCTCTCCAGGTCTCATCTCAGTCAGTGTAGGCCAGTGTATA 3619  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
305 AATGAAATGGCACAATCTCTCCAGGTCTCATCTCAGTCAGTGTAGGCCAGTGTATA 364  
QY 3620 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3679  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
365 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 424

QY 3680 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCCGATGATAAATAGTTATTTTAAG 3739  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
425 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCCGATGATAAATAGTTATTTTAAG 484  
QY 3740 AAACCTAATTCCCACTGAACCTTAAATCATCAAAAGCAGCAGTGGCCTCTAGCTTTTACTCCT 3799  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
485 AAACCTAATTCCCACTGAACCTTAAATCATCAAAAGCAGCAGTGGCCTCTAGCTTTTACTCCT 544  
QY 3800 TTGCTGAAA 3808  
Db ||||||||  
545 TTGCTGAAA 553

Search completed: February 11, 2003, 21:15:50  
Job time : 4991.35 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:31:44 ; Search time 37.6068 Seconds  
(without alignments)  
3897.581 Million cell updates/sec

Title: US-09-843-159B-3  
Perfect score: 5769  
Sequence: 1 GFGKDDVVEYLLONGASVQA.....AYPEYLITYQIMRPGCMVDG 1100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	5769	100.0	1100	22	Tankyrase homologue
2	5766	99.9	1166	22	Human tankyrase2 T
3	5766	99.9	1169	22	Human tankyrase2 I
4	5766	99.9	1169	22	Human tankyrase2 C
5	5766	99.9	1240	22	Tankyrase homologue
6	5766	99.9	1262	22	Human tankyrase2 C
7	5766	99.9	1385	22	Human tankyrase2 T
8	5750	99.7	1166	21	Human tankyrase II
9	5717	99.1	1166	22	Human tankyrase ho
10	5605	97.2	1074	20	Human Grb7 effecto

11	5201	90.2	1166	22	Human SPANK. Homo
12	4932.5	85.5	1333	21	Human tankyrase II
13	4889.5	84.8	1327	21	Human tankyrase I
14	4889.5	84.8	1327	21	Human tankyrase.
15	4889.5	84.8	1327	22	Human tankyrase. S
16	4889.5	84.8	1327	23	Human tankyrase.
17	4859	84.2	1431	21	Human tankyrase II
18	4703.5	81.5	991	22	Mouse SPANK. Mus
19	4624.5	80.2	907	22	Human breast cance
20	4102	71.1	784	22	Human tankyrase2 C
21	3959	68.6	756	22	Human tankyrase2 C
22	3835.5	66.5	1181	22	Drosophila melanog
23	3835.5	66.5	1181	22	Drosophila tankyra
24	3596	62.3	802	23	Protein of App rel
25	3093	53.6	949	21	Human truncated ta
26	3093	53.6	949	21	Human truncated ta
27	2367	41.0	522	22	Human tankyrase2 c
28	2165	37.5	415	22	Human Grb7 effecto
29	2053.5	35.6	1099	22	Human tankyrase2 e
30	1942	33.7	385	22	Human T cell leuka
31	1918.5	33.3	673	21	Human truncated ta
32	1918.5	33.3	673	23	Human TRFL TANK2 b
33	1895	32.8	362	22	Human breast cance
34	1706	29.6	331	22	Human protein sequ
35	1395	24.2	261	22	Human ORFX ORF3133
36	1251	21.7	240	21	Novel human diagno
37	875.5	15.2	4274	22	Novel human diagno
38	875.5	15.2	4386	22	Human tankyrase II
39	874	15.1	160	21	Human tankyrase I
40	874	15.1	160	22	Human tankyrase I
41	860.5	14.9	1872	22	Human secreted pro
42	860	14.9	193	22	Human tankyrase I
43	852.5	14.8	4397	22	Novel human diagno
44	836	14.5	1549	22	Drosophila melanog
45	836	14.5	1549	22	Drosophila melanog

ALIGNMENTS

RESULT 1  
AA97748  
ID AA97748 standard; Protein: 1100 AA.  
XX  
AC AA97748;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Tankyrase homologue isotype 1 protein sequence.  
XX  
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;  
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;  
KW chromosome mapping; gene therapy; vaccine.  
XX  
OS Unidentified.  
XX  
PN WO200130987-A2.  
XX  
PD 03-MAY-2001.  
XX  
PF 25-OCT-2000; 2000WO-US41528.  
XX  
PR 25-OCT-1999; 99US-0427154.  
XX  
PA (RIGE-) RIGEL PHARM INC.  
XX  
PI Luo Y, Chan E, Xu X, Huang B;  
XX  
DR WPI: 2001-300503/31.  
XX  
DR N-PSDB; AAA91487.  
XX  
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for  
PT inducing or preventing cell proliferation in cells, and for diagnosing,

treating or preventing cell cycle associated disorders such as cancer

Claim 22; Fig 3; 63pp; English.

This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein of the invention. The invention also relates to the TaHo-2 protein. The TaHo proteins are useful for inducing or preventing cell proliferation in cells, and in the study or treatment of conditions mediated by the cell cycle proteins, such as to diagnose, treat or prevent cell cycle associated disorders, preferably cancer. The TaHo coding sequences are useful as hybridisation probes, in chromosome and gene mapping and in the generation of anti-sense DNA and RNA. The coding sequences are also useful for the preparation of TaHo, for generating either transgenic animals or knock out animals which, in turn, are useful in a development and screening of therapeutically useful agents, in gene therapy, as vaccine, and for construction of hybridisation probes for mapping the gene which encodes TaHo and for the genetic analysis of individuals with genetic disorders. The TaHo proteins, and their coding sequences are useful in screening assays.

SQ Sequence 1100 AA;

Query Match 100.0%; Score 5769; DB 22; Length 1100;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGRKDVVEYLQNGASVQARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60  
DB 1 GFGRKDVVEYLQNGASVQARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60  
QY 61 PLHEAAIKGKIDVICVLLQHGAEPTIRNTDGTALDPSAKAVLGEYKDKDELLESAR 120  
DB 61 PLHEAAIKGKIDVICVLLQHGAEPTIRNTDGTALDPSAKAVLGEYKDKDELLESAR 120  
QY 121 SGNEEKMWALLTPLNVNCHASDGRKSTPLHLAGYNRVKIVQLLQHGADVHAKDKGDLV 180  
DB 121 SGNEEKMWALLTPLNVNCHASDGRKSTPLHLAGYNRVKIVQLLQHGADVHAKDKGDLV 180  
QY 181 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240  
DB 181 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240  
QY 241 NCHNSAIDLAPTQPKERLAYEFKHSLLQAAREADVTRIKKHSLEWMVNFHPQTHET 300  
DB 241 NCHNSAIDLAPTQPKERLAYEFKHSLLQAAREADVTRIKKHSLEWMVNFHPQTHET 300  
QY 301 ALHCAASPPYKPKQICELLLRKGANINEKTKFELPLHVASEKAHNDVVVVVKEAKV 360  
DB 301 ALHCAASPPYKPKQICELLLRKGANINEKTKFELPLHVASEKAHNDVVVVVKEAKV 360  
QY 361 NALDNLGOTSLHRAAYCGHLOTCRLLSYGCDDNIIISLQFTALQMGNEVQQLLEGIS 420  
DB 361 NALDNLGOTSLHRAAYCGHLOTCRLLSYGCDDNIIISLQFTALQMGNEVQQLLEGIS 420  
QY 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480  
DB 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480  
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEV 540  
DB 481 QHGADVHAKDKGGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEV 540  
QY 541 ICKLLLOHGADPTKKNRDCNTPLDLVKDGDTDIQLLRGDAALLDAARKGGLARVKKLSS 600  
DB 541 ICKLLLOHGADPTKKNRDCNTPLDLVKDGDTDIQLLRGDAALLDAARKGGLARVKKLSS 600  
QY 601 PDNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASVGHVD 660  
DB 601 PDNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASVGHVD 660  
QY 661 VAALLIKYNACVNATDKWAFTPLHEAAOKGRTOQLCALLAHGADPTLKNQEGOTPLDLVS 720  
DB 661 VAALLIKYNACVNATDKWAFTPLHEAAOKGRTOQLCALLAHGADPTLKNQEGOTPLDLVS 720

DB 661 VAALLIKYNACVNATDKWAFTPLHEAAOKGRTOQLCALLAHGADPTLKNQEGOTPLDLVS 720  
QY 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGPPSSLSAASSLDNLG 780  
DB 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGPPSSLSAASSLDNLG 780  
QY 781 SFSSELSSVVSSSGTEGASSLEKKEVPGVDFSIQTQVRNLGLEHLMDFIEREQITLDVLVE 840  
DB 781 SFSSELSSVVSSSGTEGASSLEKKEVPGVDFSIQTQVRNLGLEHLMDFIEREQITLDVLVE 840  
QY 841 MGHKELKEIGINAYCHRRKLIKGVRELISGOGLNPYLTLNTSGSSTLIDLSPDDKEFQ 900  
DB 841 MGHKELKEIGINAYCHRRKLIKGVRELISGOGLNPYLTLNTSGSSTLIDLSPDDKEFQ 900  
QY 901 SVEEMOSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEENHNHANE 960  
DB 901 SVEEMOSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEENHNHANE 960  
QY 961 RMLFHGSPFVNAILHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHKD 1020  
DB 961 RMLFHGSPFVNAILHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHKD 1020  
QY 1021 RSCVICHRLQLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYIRGEQ 1080  
DB 1021 RSCVICHRLQLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYIRGEQ 1080  
QY 1081 AYPEYLITYQIMRPEGMDVG 1100  
DB 1081 AYPEYLITYQIMRPEGMDVG 1100

# RESULT 2

AAB66295

ID AAB66295 standard; Protein; 1166 AA.

AC AAB66295;

XX XX

DT 05-APR-2001 (first entry)

XX Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.

DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
KW Inflammatory disorder.

XX Homo sapiens.

OS OS

XX WO200100849-A1.

PN 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

PF 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

PA Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI: 2001-102896/11.

DR N-PSDB; AAF63953.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by

PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,

PT inflammatory and autoimmune disorders

XX Claim 3; Page 200-203; 242pp; English.

XX The present invention provides the protein and coding sequence for the

CC human tankyrase2 protein. This is found in two different versions,

CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has

CC polyADP-ribosylation activity and is involved in the modification of

CC TRF1, which is a telomere-specific binding protein. The regulation of

CC telomere length, in which TRF1 has a role, is linked to ageing and



CC	cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.	
CC		
XX		
SQ	Sequence 1166 AA;	
	Query Match 99.9%; Score 5766; DB 22; Length 1166;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GFRKDVVEYLLQNGASVOARDGGLIPLHNACSFGEAEVNNLLRHGADPNARDNNWYT 60	
Db	67 GFRKDVVEYLLQNGANVOARDGGLIPLHNACSFGEAEVNNLLRHGADPNARDNNWYT 126	
QY	61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGR TALDADPSAKAVLTGEYKKDELLESAR 120	
Db	127 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGR TALDADPSAKAVLTGEYKKDELLESAR 186	
QY	121 SGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 180	
Db	187 SGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 246	
QY	181 PLHNACSYGHEVTELLVYKHGACVNAMDLMQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240	
Db	247 PLHNACSYGHEVTELLVYKHGACVNAMDLMQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306	
QY	241 NCHNKAIDLAPTPQKLERLAYEFKGHSLLQAAAREADVTIRIKKHL SLEMVNFKHPQTHET 300	
Db	307 NCHNKAIDLAPTPQKLERLAYEFKGHSLLQAAAREADVTIRIKKHL SLEMVNFKHPQTHET 366	
QY	301 ALHCAASPYPKFKOICE LLLRKGANINEKTFEFTPLHVASEKAHNDVVEVVKHAEKV 360	
Db	367 ALHCAASPYPKFKOICE LLLRKGANINEKTFEFTPLHVASEKAHNDVVEVVKHAEKV 426	
QY	361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIIISLQGTALQMGNEVVOQLQEGIS 420	
Db	427 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIIISLQGTALQMGNEVVOQLQEGIS 486	
QY	421 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDTEGROSTPLHFAAGYNRVSVVEYLL 480	
Db	487 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDTEGROSTPLHFAAGYNRVSVVEYLL 546	
QY	481 QHGADVHAKDKGGLVPLHNACSYGHEVAFELLVKHGAVNVNADLWKF TPLHEAAAKGKYE 540	
Db	547 QHGADVHAKDKGGLVPLHNACSYGHEVAFELLVKHGAVNVNADLWKF TPLHEAAAKGKYE 606	
QY	541 ICKLLQHGADPTKKNRDGNTPDLVKGDTDIQDLRGDAALLDAAKGCCLARVKKLS 600	
Db	607 ICKLLQHGADPTKKNRDGNTPDLVKGDTDIQDLRGDAALLDAAKGCCLARVKKLS 666	
QY	601 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 660	
Db	667 PDVNCRDTOGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 726	
QY	661 VAALLIKYNACVNATDKWAF TPLHEAAQKRTQLCALLLAHGADPTLKNQEGTPLDLVS 720	
Db	727 VAALLIKYNACVNATDKWAF TPLHEAAQKRTQLCALLLAHGADPTLKNQEGTPLDLVS 786	
QY	721 ADDYSALLTAMPSPALPSCYKQVNLNCRSPGATADALSSGSPSSLSAASLNLNSG 780	
Db	787 ADDYSALLTAMPSPALPSCYKQVNLNCRSPGATADALSSGSPSSLSAASLNLNSG 846	
QY	781 SFSSELSSVSSSGTEGASSLEKKEVPQGVDFSTIQFVFNGLGLEHLMDFEREQITLDVLVE 840	
Db	847 SFSSELSSVSSSGTEGASSLEKKEVPQGVDFSTIQFVFNGLGLEHLMDFEREQITLDVLVE 906	
QY	841 MGHEKELGIGNAYGHRHKLKIGVERLISGOQGLNPNLYTLTNTSGGTILIDLSPDQKEFQ 900	
Db	907 MGHEKELGIGNAYGHRHKLKIGVERLISGOQGLNPNLYTLTNTSGGTILIDLSPDQKEFQ 966	
QY	901 SVEEEMQSTVREHRDGGHAGGIFNRYNTLKIQVCNKKLWERYTHRKEVSEENHNHANE 960	
Db	967 SVEEEMQSTVREHRDGGHAGGIFNRYNTLKIQVCNKKLWERYTHRKEVSEENHNHANE 1026	

QY	961 RMLFHGSPFYNAIIHKGFDERHAYIGGMFCAGIYFAENSSKSNQYVYVIGGTCGPVHKD 1020	
Db	1027 RMLFHGSPFYNAIIHKGFDERHAYIGGMFCAGIYFAENSSKSNQYVYVIGGTCGPVHKD 1086	
QY	1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVYIRGEQ 1080	
Db	1087 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVYIRGEQ 1146	
QY	1081 AYPEYLITYQIMRPEGMVDG 1100	
Db	1147 AYPEYLITYQIMRPEGMVDG 1166	
	RESULT 3	
ID	AA66278	
XX	AA66278 standard; Protein: 1169 AA.	
AC	AA66278;	
XX	AC	
XX	05-APR-2001 (first entry)	
DT	XX	
XX	Human tankyrase2 related protein sequence SEQ ID NO: 2.	
DE	XX	
XX	Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging; inflammatory disorder.	
KW	XX	
XX	Homo sapiens.	
OS	XX	
XX	WO200100849-A1.	
PN	XX	
PD	XX	
XX	04-JAN-2001.	
PF	XX	
XX	28-JUN-2000; 2000WO-US17827.	
PR	XX	
XX	29-JUN-1999; 99US-0141582.	
PA	XX	
XX	(ICOS-) ICOS CORP.	
PI	XX	
XX	Christenson E, Demaggio AJ, Goldman PS, McElligott DL;	
DR	XX	
XX	WPI; 2001-102896/11.	
DR	XX	
XX	N-PSDB; AAF63837.	
PT	XX	
XX	New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -	
PS	XX	
XX	Disclosure: Page 109-113; 242pp; English.	
CC	XX	
XX	The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has poly(ADP-ribose)ylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.	
SQ	XX	
	Sequence 1169 AA;	
	Query Match 99.9%; Score 5766; DB 22; Length 1169;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GFRKDVVEYLLQNGASVOARDGGLIPLHNACSFGEAEVNNLLRHGADPNARDNNWYT 60	
Db	70 GFRKDVVEYLLQNGANVOARDGGLIPLHNACSFGEAEVNNLLRHGADPNARDNNWYT 129	
QY	61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGR TALDADPSAKAVLTGEYKKDELLESAR 120	
Db	130 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGR TALDADPSAKAVLTGEYKKDELLESAR 189	
QY	121 SGNEEKMMALLPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 180	

```
Db 190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 249
QY 181 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 240
Db 250 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 309
QY 241 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 300
Db 310 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 369
QY 301 ALHCAASPYPRKQICELLRLKGANINEKTKFETPLHVASEKAHNDVVEVVVKHEAKV 360
Db 370 ALHCAASPYPRKQICELLRLKGANINEKTKFETPLHVASEKAHNDVVEVVVKHEAKV 429
QY 361 NALDNLGQTSLHRAAYCGHLOTCRLLISYGGDPNIIISLOGFTALQMGNEVQQLQEGIS 420
Db 430 NALDNLGQTSLHRAAYCGHLOTCRLLISYGGDPNIIISLOGFTALQMGNEVQQLQEGIS 489
QY 421 LGNSEADROLLEAAKAGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
Db 490 LGNSEADROLLEAAKAGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 549
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGAVNVADUMKFTPLHEAAAKGYE 540
Db 550 QHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGAVNVADUMKFTPLHEAAAKGYE 609
QY 541 ICKLLQHGADPTKKNRDNPTPLDLVKDGDPTDIDLLRGDAALLDAKKGCLARVKKLSS 600
Db 610 ICKLLQHGADPTKKNRDNPTPLDLVKDGDPTDIDLLRGDAALLDAKKGCLARVKKLSS 669
QY 601 PDVNCRTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 660
Db 670 PDVNCRTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 729
QY 661 VAALLIKYNACVNATDKWAFPLHEAAKQRTQCALLLAHGADPTLKNQEGQTPDLVLS 720
Db 730 VAALLIKYNACVNATDKWAFPLHEAAKQRTQCALLLAHGADPTLKNQEGQTPDLVLS 789
QY 721 ADDVSALLTAMPSPALPSCYKPOVLNVRSPGATADALSSGSPSPSSLSAASLDNLG 780
Db 790 ADDVSALLTAMPSPALPSCYKPOVLNVRSPGATADALSSGSPSPSSLSAASLDNLG 849
QY 781 SFSLSWVSSSGTEGASSLEKKEVPVGSITQFVRLNGLHLMDFEREQITDLVIVE 840
Db 850 SFSLSWVSSSGTEGASSLEKKEVPVGSITQFVRLNGLHLMDFEREQITDLVIVE 909
QY 841 MGKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
Db 910 MGKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 969
QY 901 SVEEEMQSTVREHRDGGHAGIFNRYNLIKQVCNKKLWERYTHRKEYVSEENHNHANE 1029
Db 970 SVEEEMQSTVREHRDGGHAGIFNRYNLIKQVCNKKLWERYTHRKEYVSEENHNHANE 1099
QY 961 RMLFHGSPFFNAILHKGFDERHAYIGMFGAGIYFAENSCKSNQYVIGGGTGCPVHKD 1020
Db 1030 RMLFHGSPFFNAILHKGFDERHAYIGMFGAGIYFAENSCKSNQYVIGGGTGCPVHKD 1089
QY 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHSHSVTGRPSVNGLALAEYVIRGEQ 1080
Db 1090 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHSHSVTGRPSVNGLALAEYVIRGEQ 1149
QY 1081 AYPEYLITQIMRPEGMVDG 1100
Db 1150 AYPEYLITQIMRPEGMVDG 1169
```

```
RESULT 4
ID AAB66288
XX AAB66288 standard; Protein; 1169 AA.
AC AAB66288;
```

```
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
KW Human: tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17827.
PR 29-JUN-1999; 99US-0141582.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
DR WPI; 2001-102896/11.
XX
DR N-PSDB; AAF63926.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX inflammatory and autoimmune disorders.
XX
PS Example 1; Page 162-1665; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to ageing and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX
SQ Sequence 1169 AA;
```

```
Query Match 99.9%; Score 5766; DB 22; Length 1169;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GFGRKDWVEYLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNY 60
Db 70 GFGRKDWVEYLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNY 129
QY 61 PLHEAAIKGIDYCVILVLOHGAETIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 120
Db 130 PLHEAAIKGIDYCVILVLOHGAETIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 189
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 180
Db 190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 249
QY 181 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 240
Db 250 PLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 309
QY 241 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 300
Db 310 NCHNKSALDAPTLPOLKERLAYEFKGHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHET 369
QY 301 ALHCAASPYPRKQICELLRLKGANINEKTKFETPLHVASEKAHNDVVEVVVKHEAKV 360
Db 370 ALHCAASPYPRKQICELLRLKGANINEKTKFETPLHVASEKAHNDVVEVVVKHEAKV 429
QY 361 NALDNLGQTSLHRAAYCGHLOTCRLLISYGGDPNIIISLOGFTALQMGNEVQQLQEGIS 420
Db 430 NALDNLGQTSLHRAAYCGHLOTCRLLISYGGDPNIIISLOGFTALQMGNEVQQLQEGIS 489
```

```
QY 421 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
Db 490 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 549
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 540
Db 550 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 609
QY 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGDPTDIDQLRGGDAALDAKKGCLARVKKLSS 600
Db 610 ICKLLQHGADPTKKNRDGNTPLDLVKDGDPTDIDQLRGGDAALDAKKGCLARVKKLSS 669
QY 601 PDNVNCRDTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 660
Db 670 PDNVNCRDTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVD 729
QY 661 VAALLIKYNACVNATDKWAFPLHEAAQKGRTOCALLLAHGADPTLKNQEGQTPDLVS 720
Db 730 VAALLIKYNACVNATDKWAFPLHEAAQKGRTOCALLLAHGADPTLKNQEGQTPDLVS 789
QY 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSPSLSAASLDNLG 780
Db 790 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSPSLSAASLDNLG 849
QY 781 SFSELSVVSSSGTEGASLEKKEVPGVDFESTQFVRNLGLEHLMDFEREQITLDLVE 840
Db 850 SFSELSVVSSSGTEGASLEKKEVPGVDFESTQFVRNLGLEHLMDFEREQITLDLVE 909
QY 841 MGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTNTSGSGTILIDLSPDDKEFQ 900
Db 910 MGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTNTSGSGTILIDLSPDDKEFQ 969
QY 901 SVEEEMQSVRHRDGGHAGGTIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNHANE 960
Db 970 SVEEEMQSVRHRDGGHAGGTIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNHANE 1029
QY 961 RMLFHGSPFVNALIHKGFDERHAYIGGMFGAGIYFAENSSKNQYVYIGGTCGCPVHKD 1020
Db 1030 RMLFHGSPFVNALIHKGFDERHAYIGGMFGAGIYFAENSSKNQYVYIGGTCGCPVHKD 1089
QY 1021 RSCYICHRQLLCFVRTVLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
Db 1090 RSCYICHRQLLCFVRTVLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1149
QY 1081 AYPEYLITYQIMRPEGWVDG 1100
Db 1150 AYPEYLITYQIMRPEGWVDG 1169

RESULT 5
AA97749
XX AAY97749 standard; Protein; 1240 AA.
XX
XX AAY97749;
XX
XX 06-AUG-2001 (first entry)
XX
XX Tankyrase homologue isotype 2 protein sequence.
XX
XX Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
XX cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
XX chromosome mapping; gene therapy; vaccine.
XX
XX Unidentified.
XX
XX W0200130987-A2.
XX
XX 03-MAY-2001.
XX
XX 25-OCT-2000; 2000WO-US41528.
XX
XX 25-OCT-1999; 99US-0427154.
XX
PR
```

```
XX (RIGE-) RIGEL PHARM INC.
PA Luo Y, Chan E, Xu X, Huang B;
PI WPI; 2001-300503/31.
XX N-PSDB; AAA91488.
DR Novel recombinant cell cycle polypeptide, tankyrase H useful for
PT inducing or preventing cell proliferation in cells, and for diagnosing,
PT treating or preventing cell cycle associated disorders such as cancer
PT
XX Claim 22; Fig 4; 63pp; English.
PS
XX This sequence is the Tankyrase homologue isotype 2 (TaHo-2) protein
XX of the invention. The invention also relates to the TaHo-2 protein.
CC The TaHo proteins are useful for inducing or preventing cell
CC proliferation in cells, and in the study or treatment of conditions
CC mediated by the cell cycle proteins, such as to diagnose, treat or
CC prevent cell cycle associated disorders, preferably cancer. The TaHo
CC coding sequences are useful as hybridisation probes, in chromosome and
CC gene mapping and in the generation of anti-sense DNA and RNA. The coding
CC sequences are also useful for the preparation of TaHo, for generating
CC either transgenic animals or knock out animals which, in turn, are useful
CC in a development and screening of therapeutically useful agents, in gene
CC therapy, as vaccine, and for construction of hybridisation probes for
CC mapping the gene which encodes TaHo and for the genetic analysis of
CC individuals with genetic disorders. The TaHo proteins, and their coding
CC sequences are useful in screening assays.
XX
XX Sequence 1240 AA:
Query Match 99.9%; Score 5766; DB 22; Length 1240;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GFGRKDVVEYLLONGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYNT 60
Db 141 GFGRKDVVEYLLONGANVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYNT 200
QY 61 PLHEAAATGKIDVCIVLLOHGAETPIRNTDGTALDLPADPSAKAVLTGEYKKDELLESAR 120
Db 201 PLHEAAATGKIDVCIVLLOHGAETPIRNTDGTALDLPADPSAKAVLTGEYKKDELLESAR 260
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 180
Db 261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 320
QY 181 PLHNACSYGHEVTELLVKGACYNAMDWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
Db 321 PLHNACSYGHEVTELLVKGACYNAMDWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
QY 241 NCHNKSATDLAPTOLKERLAYEFKGSLLQAAAREADVTRIKKHLSEMVNFKHPQTHET 300
Db 381 NCHNKSATDLAPTOLKERLAYEFKGSLLQAAAREADVTRIKKHLSEMVNFKHPQTHET 440
QY 301 ALHCAASAPYKPKQICELLRLKGANINEKTEFPLTHVASEKAHNDVVEVVKHEAVK 360
Db 441 ALHCAASAPYKPKQICELLRLKGANINEKTEFPLTHVASEKAHNDVVEVVKHEAVK 500
QY 361 NALDNLGOTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOGFTALQMGNNVQQLQEGIS 420
Db 501 NALDNLGOTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOGFTALQMGNNVQQLQEGIS 560
QY 421 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480
Db 561 LGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 620
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 540
Db 621 QHGADVHAKDKGGLVPLHNACSYGHEVAELLVKGAVNVADLWKFTPLHEAAAKGYE 680
```

QY 541 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDIDLLRGDAALLDAKKGGLARVKLSS 600  
 Db 681 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDIDLLRGDAALLDAKKGGLARVKLSS 740  
 QY 601 PDVNCRDTCGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660  
 Db 741 PDVNCRDTCGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800  
 QY 661 VAALLIKYNACVNATDKWAFPLHEAAKGRTOQLCALLLAHGAOPTLKNQSGOTPLDLVS 720  
 Db 801 VAALLIKYNACVNATDKWAFPLHEAAKGRTOQLCALLLAHGAOPTLKNQSGOTPLDLVS 860  
 QY 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 780  
 Db 861 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 920  
 QY 781 SFSLSVSSVSSGTEGASSLEKKEVPGVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 840  
 Db 921 SFSLSVSSVSSGTEGASSLEKKEVPGVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 980  
 QY 841 MGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYTLTNTSGSGTILIDLSPDDKEFQ 900  
 Db 981 MGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPYTLTNTSGSGTILIDLSPDDKEFQ 1040  
 QY 901 SVEEMOSTVREHRDGGHAGGIFNRYNLIKQVCNKKLWERTHRRKEVSEENHNHANE 960  
 Db 1041 SVEEMOSTVREHRDGGHAGGIFNRYNLIKQVCNKKLWERTHRRKEVSEENHNHANE 1100  
 QY 961 RMLFHGSPFFNAIHKGFDERHAYIGMFCAGIYFAENSSKSNQYVYIGGGTGPCPVHKD 1020  
 Db 1101 RMLFHGSPFFNAIHKGFDERHAYIGMFCAGIYFAENSSKSNQYVYIGGGTGPCPVHKD 1160  
 QY 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSHVTRGPSVNGLALAEYVYIRGEQ 1080  
 Db 1161 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSHVTRGPSVNGLALAEYVYIRGEQ 1220  
 QY 1081 AYPEYLITYQIMRPEGMVDG 1100  
 Db 1221 AYPEYLITYQIMRPEGMVDG 1240  
 RESULT 6  
 AAB66290  
 ID AAB66290 standard; Protein; 1262 AA.  
 AC AAB66290;  
 XX  
 DT 05-APR-2001 (first entry)  
 XX  
 DE Human tankyrase2 clone consensus protein SEQ ID NO: 107.  
 XX  
 KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasia; aging;  
 KW inflammatory disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200100849-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 28-JUN-2000; 2000WO-US17827.  
 XX  
 PR 29-JUN-1999; 99US-0141582.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 XX  
 DR WPI; 2001-102896/11.  
 XX  
 DR N-PSDB; AAF63930.  
 XX  
 PT New tankyrase2 polypeptides, useful for treating conditions mediated by  
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,

PT inflammatory and autoimmune disorders -  
 XX  
 PS Example 2; Page 173-176; 242pp; English.  
 XX

The present invention provides the protein and coding sequence for the  
 human tankyrase2 protein. This is found in two different versions,  
 designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
 polyADP-ribosylation activity and is involved in the modification of  
 TRF1, which is a telomere-specific binding protein. The regulation of  
 telomere length, in which TRF1 has a role, is linked to ageing and  
 cancer. The sequences are useful in the treatment of cancers and  
 inflammatory disorders.

Sequence 1262 AA:

Query Match 99.9%; Score 5766; DB 22; Length 1262;  
 Best local Similarity 99.9%; Pred. No. 0;  
 Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGEKDVVEYLLQNGASVQARDGGLIPLHNACSGFHAENVNLLRHGADPNARDNNYNT 60  
 Db 163 GFGEKDVVEYLLQNGANVQARDGGLIPLHNACSGFHAENVNLLRHGADPNARDNNYNT 222  
 QY 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 120  
 Db 223 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESAR 282  
 QY 121 SGNEEKMAALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 180  
 Db 283 SGNEEKMAALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 342  
 QY 181 PLHNACSYGHEYTELLVKGACVNMALQOFTPLHEAASKNRVEVCSSLISYGADPTLL 240  
 Db 343 PLHNACSYGHEYTELLVKGACVNMALQOFTPLHEAASKNRVEVCSSLISYGADPTLL 402  
 QY 241 NCHNKSAIDLAPTQPKERLAYEFKGHSLLOAAREADVTRIKHLSLEMYNFKHPQTHET 300  
 Db 403 NCHNKSAIDLAPTQPKERLAYEFKGHSLLOAAREADVTRIKHLSLEMYNFKHPQTHET 462  
 QY 301 ALHCAAAAPYPRKQICEILLRKGANINEKTEFLPLHVASEKAHNDVYVVKHEAV 360  
 Db 463 ALHCAAAAPYPRKQICEILLRKGANINEKTEFLPLHVASEKAHNDVYVVKHEAV 522  
 QY 361 NALDNLGOTSLEHRAAYCGHLOTCRLLLSYGCDPNIIISLOGFTALOMGNENVOQLLEGIS 420  
 Db 523 NALDNLGOTSLEHRAAYCGHLOTCRLLLSYGCDPNIIISLOGFTALOMGNENVOQLLEGIS 582  
 QY 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRYSVVEYLL 480  
 Db 583 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRYSVVEYLL 642  
 QY 481 QHGADVHAKDKGGLVPLHNACSYGHEYTELLVKGACVNMALQOFTPLHEAASKNRVEVCSSLISYGADPTLL 540  
 Db 643 QHGADVHAKDKGGLVPLHNACSYGHEYTELLVKGACVNMALQOFTPLHEAASKNRVEVCSSLISYGADPTLL 600  
 QY 541 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDIDLLRGDAALLDAKKGGLARVKLSS 762  
 Db 703 ICKLLQHGADPTKKNRDNPTPLDLVKDGDITDIDLLRGDAALLDAKKGGLARVKLSS 822  
 QY 601 PDVNCRDTCGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660  
 Db 763 PDVNCRDTCGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 822  
 QY 661 VAALLIKYNACVNATDKWAFPLHEAAKGRTOQLCALLLAHGAOPTLKNQSGOTPLDLVS 720  
 Db 823 VAALLIKYNACVNATDKWAFPLHEAAKGRTOQLCALLLAHGAOPTLKNQSGOTPLDLVS 882  
 QY 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 780  
 Db 883 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASLDNLG 942  
 QY 781 SFSLSVSSVSSGTEGASSLEKKEVPGVDFSTITQFVRNLGLEHLMDFEREQITDLVLE 840

Db 943 SFSELSSVSSSTEGASSLEKKEVPVDFSTQFVRNLGLEHLMDFEREQITLDVLVE 1002  
Qy 841 MGKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTLNTSGSTILIDLSPDDKEFQ 900  
Db 1003 MGKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTLNTSGSTILIDLSPDDKEFQ 1062  
Qy 901 SVEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 960  
Db 1063 SVEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 1122  
Qy 961 RMLFHGSPFNALIIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGTGCPVHKD 1020  
Db 1123 RMLFHGSPFNALIIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGTGCPVHKD 1182  
Qy 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080  
Db 1183 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVIYRGEQ 1242  
Qy 1081 AYPEYLITYQIMRPEGMVDG 1100  
Db 1243 AYPEYLITYQIMRPEGMVDG 1262

RESULT 7  
AAB66294  
ID AAB66294 standard; Protein; 1385 AA.

XX AC AAB66294;  
XX 05-APR-2001 (first entry)  
XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.  
XX DE  
XX KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
XX KW inflammatory disorder.

XX OS Homo sapiens.  
XX PN WO200100849-A1.  
XX PD 04-JAN-2001.  
XX PF 28-JUN-2000; 2000WO-US17827.  
XX PR 29-JUN-1999; 99US-0141582.  
XX PA (ICOS-) ICOS CORP.  
XX PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
XX WPI: 2001-102896/11.  
XX DR N-PSDB; AAF63952.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
XX poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
XX inflammatory and autoimmune disorders.  
XX Claim 2: Page 191-194; 242pp; English.

XX The present invention provides the protein and coding sequence for the  
XX human tankyrase2 protein. This is found in two different versions,  
XX designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
XX poly(ADP-ribose)ylation activity and is involved in the modification of  
XX TRF1, which is a telomere-specific binding protein. The regulation of  
XX telomere length, in which TRF1 has a role, is linked to ageing and  
XX cancer. The sequences are useful in the treatment of cancers and  
XX inflammatory disorders.

XX Sequence 1385 AA;

Query Match 99.9%; Score 5766; DB 22; Length 1385;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGRKDVVEYLLONGCASVOARDGGLIPLHNACSFGEAEVNVNLLRHGADPNARDNNYNT 60  
Db 286 GFGRKDVVEYLLONGANVOARDGGLIPLHNACSFGEAEVNVNLLRHGADPNARDNNYNT 345  
Qy 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKDELLESAR 120  
Db 346 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKDELLESAR 405  
Qy 121 SNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLOHGADVHAKDKGLV 180  
Db 406 SNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLOHGADVHAKDKGLV 465  
Qy 181 PLHNACSYGHEVETELLVKHGACVNAMDLOFTPLHEAASKNRVEVCSSLISYGADPTLL 240  
Db 466 PLHNACSYGHEVETELLVKHGACVNAMDLOFTPLHEAASKNRVEVCSSLISYGADPTLL 525  
Qy 241 NCHNKSALDAPTPQIKERLAYEFKGHSLLQAARADVTIRIKKHLISLWVNFKHPQTHET 300  
Db 526 NCHNKSALDAPTPQIKERLAYEFKGHSLLQAARADVTIRIKKHLISLWVNFKHPQTHET 585  
Qy 301 ALHCAAAAPYPKRKQICELLRLKGANINEKTEFTPLHVASEKAHNDVVEVVKHEAKV 360  
Db 586 ALHCAAAAPYPKRKQICELLRLKGANINEKTEFTPLHVASEKAHNDVVEVVKHEAKV 645  
Qy 361 NALDNLGOTSLHRAAYCGHLQTCRLLLSYCCDPNIIISLQFTALQMGNEVVOQLQEGIS 420  
Db 646 NALDNLGOTSLHRAAYCGHLQTCRLLLSYCCDPNIIISLQFTALQMGNEVVOQLQEGIS 705  
Qy 421 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480  
Db 706 LGNSEADROLLEAAKAGDVETVKKLTCTQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 765  
Qy 481 QHGADVHAKDKGLVPLHNACSYGHEVAELLVKHGAVNVADLWKFPTPLHEAAAKGKYE 540  
Db 766 QHGADVHAKDKGLVPLHNACSYGHEVAELLVKHGAVNVADLWKFPTPLHEAAAKGKYE 825  
Qy 541 ICKLLLOHGADPTKKNRDGNTPLDLVKDGDITQDILLRDAALLDAKKGCLARVKKLS 600  
Db 826 ICKLLLOHGADPTKKNRDGNTPLDLVKDGDITQDILLRDAALLDAKKGCLARVKKLS 885  
Qy 601 PDVNCRDGTGRHSTPLHLAAGYNVLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 660  
Db 886 PDVNCRDGTGRHSTPLHLAAGYNVLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 945  
Qy 661 VAALLIKYNACVNATDKWAFPTPLHEAAQKGTQLCALLAHGADPTLKNQEGOTPLDLVS 720  
Db 946 VAALLIKYNACVNATDKWAFPTPLHEAAQKGTQLCALLAHGADPTLKNQEGOTPLDLVS 1005  
Qy 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGPSSPSLSAASLDNLG 780  
Db 1006 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGPSSPSLSAASLDNLG 1065  
Qy 781 SFSSELSSVSSSTEGASSLEKKEVPVDFSTQFVRNLGLEHLMDFEREQITLDVLVE 840  
Db 1066 SFSSELSSVSSSTEGASSLEKKEVPVDFSTQFVRNLGLEHLMDFEREQITLDVLVE 1125  
Qy 841 MGKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTLNTSGSTILIDLSPDDKEFQ 900  
Db 1126 MGKELKEIGINAYGHRHKLKIGVERLISGOOGLNPYLTLNTSGSTILIDLSPDDKEFQ 1185  
Qy 901 SVEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 960  
Db 1186 SVEEMOSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRKEVSEENHNHANE 1245  
Qy 961 RMLFHGSPFNALIIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGTGCPVHKD 1020  
Db 1246 RMLFHGSPFNALIIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGTGCPVHKD 1305  
Qy 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080  
Db 1306 RSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVIYRGEQ 1365



02-MAY-2001 (first entry)  
Human tankyrase homolog protein (THP).  
XX  
DE Human tankyrase homolog protein; THP; gene therapy; cancer;  
KW tumour; basal cell carcinoma; therapy; genetic mapping.  
KW cytosolic.  
XX  
OS Homo sapiens.  
XX

Key Location/Qualifiers  
FT Misc-difference 368 /label= Met, Leu, Val  
FT Misc-difference 392 /label= Asn, Thr  
FT Misc-difference 415 /label= Val, Ile  
FT Misc-difference 558 /label= Gly, Glu  
FT Misc-difference 559 /label= Gly, Ala  
FT Misc-difference 764 /label= Leu, Val  
FT Misc-difference 884 /label= Asn, His, Asp, Tyr  
XX  
XX WO200104326-A1.  
XX  
XX 18-JAN-2001.  
XX  
XX 03-JUL-2000; 2000WO-EP06609.  
XX  
XX 09-JUL-1999; 99US-0350982.  
XX  
XX (PAAA ) PHARMACIA & UPJOHN SPA.  
XX  
XX Berthelsen J, Toma S, Isacchi A;  
XX  
XX WPI: 2001-168422/17.  
XX  
XX N-PSDB: AAD02578.

New tankyrase homolog protein (THP) polynucleotide and polypeptide  
useful in gene therapy, diagnosis and treatment or prevention of  
unregulated cell growth, such as cancer or tumor cell growth -  
Claim 20; Page 50-54; 60pp; English.

The present sequence is human tankyrase homolog  
protein (THP). The THP polypeptides and polynucleotides of the  
invention are useful in gene therapy and for treating or preventing  
unregulated cell growth such as cancer or tumour (e.g. basal cell  
carcinoma). The nucleic acid molecules of the invention and their  
fragments are useful for restriction fragment length polymorphism  
(RFLP) associated with certain disorders, as well as for genetic  
mapping. Antisense oligonucleotides, or fragments of nucleic acid  
encoding THP are useful as diagnostic tools for probing the  
expression of THP gene in various tissues. THP can be used as  
antigens for raising antibodies against them and in assays for  
identifying compounds that modulate their activity. They are used in  
the manufacture of a medicament directed towards cancers or tumours.  
THP are also useful for screening compounds in a variety of drugs  
screening techniques and as a research tool for identification,  
characterisation and purification of interacting, regulatory proteins.

Sequence 1166 AA:  
Query Match 99.1%; Score 5717; DB 22; Length 1166;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1091; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
1 GFGKQWVEYLLONGASVQARDGGLIPLHNACSFSGHAEVNNLLRRHGADPNARDNNVT 60  
67 GFGKQWVEYLLONGASVQARDGGLIPLHNACSFSGHAEVNNLLRRHGADPNARDNNVT 126

Qy 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDPSAKAVLTGEYKKDELLESAR 120  
Db 127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDPSAKAVLTGEYKKDELLESAR 186  
Qy 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 180  
Db 187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLV 246  
Qy 181 PLHNACSYGHEVTELLVKHGACYNAMDQWPTPLHEAASKNRVEVCSLLISYGADPTLL 240  
Db 247 PLHNACSYGHEVTELLVKHGACYNAMDQWPTPLHEAASKNRVEVCSLLISYGADPTLL 306  
Qy 241 NCHNKSADLAPTQPKERLAYEFKGHSLLOAAREADYTRIKKHLSLEMFNFKHPQTHET 300  
Db 307 NCHNKSADLAPTQPKERLAYEFKGHSLLOAAREADYTRIKKHLSLEMFNFKHPQTHET 366  
Qy 301 ALHCAASPYPKRKQICELLRKGANINEKTEFTPLHVASEKAHNDVVEVVKHEAKV 360  
Db 367 AXHCAASPYPKRKQICELLRKGANINEKTEFTPLHVASEKAHNDVVEVVKHEAKV 426  
Qy 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOFTALQMGNEVQOLLQEGIS 420  
Db 427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOFTALQMGNEVQOLLQEGIS 486  
Qy 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480  
Db 487 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546  
Qy 481 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKHGACYNAMDQWPTPLHEAANKGYE 540  
Db 547 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKHGACYNAMDQWPTPLHEAANKGYE 606  
Qy 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDILLRGDAALLDAKKGCLARVKLSS 600  
Db 607 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDILLRGDAALLDAKKGCLARVKLSS 666  
Qy 601 PDVNCRTDQGRHSTPLHLAAGYNRVKIVOLLQHGADVNAQDKGGLIPLHNAASYGHVD 660  
Db 667 PDVNCRTDQGRHSTPLHLAAGYNRVKIVOLLQHGADVNAQDKGGLIPLHNAASYGHVD 726  
Qy 661 VAALLIKYNACVNATDKWFTPLHEAAKQKRTOLCALLAHGADPTLKNOEGOTPLDLVS 720  
Db 727 VAALLIKYNACVNATDKWFTPLHEAAKQKRTOLCALLAHGADPTLKNOEGOTPLDLVS 786  
Qy 721 ADDVSALLTAAMPSPALPCYKQVLNVRSPGATADALSSGSPSSSLSAASLDNLG 780  
Db 787 ADDVSALLTAAMPSPALPCYKQVLNVRSPGATADALSSGSPSSSLSAASLDNLG 846  
Qy 781 SFSELSVSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLE 840  
Db 847 SFSELSVSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLE 906  
Qy 841 MGHKELKEIGNAYGHRHKLKGVRLISGQGLNPVLTNTSGSGTILIDLSPPDKKEFQ 900  
Db 907 MGHKELKEIGNAYGHRHKLKGVRLISGQGLNPVLTNTSGSGTILIDLSPPDKKEFQ 966  
Qy 901 SVEEEMQSTVREHRDGGHAGGIFNRYNLTOKVCNKKLWERYTHRRKEVSEENHNANE 960  
Db 967 SVEEEMQSTVREHRDGGHAGGIFNRYNLTOKVCNKKLWERYTHRRKEVSEENHNANE 1026  
Qy 961 RMLFHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGTCGPVHKD 1020  
Db 1027 RMLFHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGTCGPVHKD 1086  
Qy 1021 RSCYICHRQLLFCRVTLGKSFLOFSAMKWAHSPGHHSVTGRPSVNGLALAEYVYRGEQ 1080  
Db 1087 RSCYICHRQLLFCRVTLGKSFLOFSAMKWAHSPGHHSVTGRPSVNGLALAEYVYRGEQ 1146  
Qy 1081 AYPEYLITYQIMRPEGMVDG 1100  
Db 1147 AYPEYLITYQIMRPEGMVDG 1166

```

RESULT 10
RAY05734
ID AA05734 standard; Protein; 1074 AA.
XX
AC AA05734;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human Grb7 effector 2.2412 protein.
XX
KW Grb7 effector; 2.2412 protein; human; signal transduction;
KW tumour marker; breast cancer; prostate cancer; prognosis;
KW diagnosis.
XX
OS Homo sapiens.
XX
PN W09915647-A1.
XX
PD 01-APR-1999.
XX
PF 23-SEP-1998; 98WO-AU00795.
XX
PR 23-SEP-1997; 97AU-0009388.
XX
PA (GARV-) GARVAN INST MEDICAL RES.
XX
PI Daly RJ, Sutherland RL;
XX
DR WPI; 1999-254707/21.
XX
DR N-PSDB; AAX25366.
XX
PT New candidate effector for the Grb7 family of signaling proteins,
PT and specific antibody, useful for detection and treatment of cancer
XX
PS Claim 9; Fig 1; 24pp; English.
XX
CC The present sequence represents a novel candidate effector for
CC the Grb7 family of signalling proteins, termed 2.2412. The
CC sequence is predicted from a partial cDNA (see AAX25366). 2.2412
CC mRNA transcripts (7 kb) were detected in all tissues examined
CC with the exception of kidney. Expression was particularly high in
CC skeletal muscle and placenta. Analysis of the sequence revealed
CC significant homology to a large number of proteins containing
CC ankyrin-like repeats. The 2.2412 gene was localised to between
CC chromosome 10q23.2 and proximal 10q23.32. Deletions in the
CC 10q22-25 region have been detected in human breast, prostate,
CC renal, small cell lung and endometrial carcinomas, glioblastoma
CC multiforme, melanoma and meningiomas. Detection of the protein
CC encoded by the 2.2412 cDNA in a sample should provide a useful
CC tumour marker and/or prognostic indicator for certain human
CC cancers, in particular breast cancer and prostate cancer.
CC Antagonism of the interaction between Grb7 family members and the
CC encoded protein should provide a novel treatment strategy for human
CC diseases exhibiting aberrant receptor tyrosine kinase signalling,
CC such as cancer. Anti-Grb7 antibodies can be used in methods of
CC detecting the presence of 2.2412 protein in a sample.
XX
SQ Sequence 1074 AA;

Query Match 97.2%; Score 5605; DB 20; Length 1074;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1069; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 27 IPLHACSGFHAENVNLLRHGADPNARDNNWYTPLHAAIKGKIDVCTVLQHGAEPTI 86
DB 1 IPLHACSGFHAENVNLLRHGADPNARDNNWYTPLHAAIKGKIDVCTVLQHGAEPTI 60
QY 87 RNTDGRALDLADPSAKAVLTGEYKKDELLESARSGNEKMMALLTPLNVNCHASDGRKS 146
DB 61 RNTDGRALDLADPSAKAVLTGEYKKDELLESARSGNEKMMALLTPLNVNCHASDGRKS 120
QY 147 TPLHLAGYNNRVKIVQLLQHGADVHAKDGLVPLHNAACSYGHEVTELLYKHGACVNA 206

```

```

Db 121 TPLHLAGYNNRVKIVQLLQHGADVHAKDGLVPLHNAACSYGHEVTELLYKHGACVNA 180
QY 207 MDLWQFTPLHAAAKNRVEVCSLLLSYGADPTLLCHNKNKSAIDLAPTQPKERLAYEFGK 266
Db 181 MDLWQFTPLHAAAKNRVEVCSLLLSYGADPTLLCHNKNKSAIDLAPTQPKERLAYEFGK 240
QY 267 HSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAAAPYPRKQICELLRLKGAN 326
Db 241 HSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAAAPYPRKQICELLRLKGAN 300
QY 327 INEKTKEFTPLHVAASEKAHNDVVEVVKHAKVNALNDLGOTSLHRAAYCGHLOTCLRL 386
Db 301 INEKTKEFTPLHVAASEKAHNDVVEVVKHAKVNALNDLGOTSLHRAAYCGHLOTCLRL 360
QY 387 LSYGCDPNIIISLOGFTALOMGNENVOQLLQEGISLGNSEADROLLEAAKAGDVETVKKLC 446
Db 361 LSYGCDPNIIISLOGFTALOMGNENVOQLLQEGISLGNSEADROLLEAAKAGDVETVKKLC 420
QY 447 TVQSVNCRDTEGROSTPLHFAAGYNRVSVVEYLLOHGADVHAKDGGGLVPLHNAACSYGHY 506
Db 421 TVQSVNCRDTEGROSTPLHFAAGYNRVSVVEYLLOHGADVHAKDGGGLVPLHNAACSYGHY 480
QY 507 EVAELLVKHGAVNVADLWKFTPLHAAAKGKYETCKILLLOHGADPTKKNRDGNTPDLV 566
Db 481 EVAELLVKHGAVNVADLWKFTPLHAAAKGKYETCKILLLOHGADPTKKNRDGNTPDLV 540
QY 567 KDGTDIDQLLLRGDAALLDAAKGCLARVKLSSPDNVNCRDQGRHSTPLHLAGYNNL 626
Db 541 KDGTDIDQLLLRGDAALLDAAKGCLARVKLSSPDNVNCRDQGRHSTPLHLAGYNNL 600
QY 627 EVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHEA 686
Db 601 EVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFPLHEA 660
QY 687 AQKGRITQCALLLAHGADPTLKNOEGOTPLDLSADVDVSAALLTAAMPSPALPCKYKQVL 746
Db 661 AQKGRITQCALLLAHGADPTLKNOEGOTPLDLSADVDVSAALLTAAMPSPALPCKYKQVL 720
QY 747 NGVRSFGATADALSSGSPSSPSAASSLNLGSGFSELSSVSSSGTEGASSLEKEVP 806
Db 721 NGVRSFGATADALSSGSPSSPSAASSLNLGSGFSELSSVSSSGTEGASSLEKEVP 780
QY 807 GVDFTSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVER 866
Db 781 GVDFTSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVER 840
QY 867 LISGQOGLNPYLTINTSGSGTILIDLSPDDKEFQSEEMQSTVREHROGHHAGGIFNRY 926
Db 841 LISGQOGLNPYLTINTSGSGTILIDLSPDDKEFQSEEMQSTVREHROGHHAGGIFNRY 900
QY 927 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNALIHKGFDERHAYIG 986
Db 901 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNALIHKGFDERHAYIG 960
QY 987 GMFAGIYFAENSCKSNQYVYIGGTCGCPVHKDRSCYICHROLLFCRVTLGKSFQFSA 1046
Db 961 GMFAGIYFAENSCKSNQYVYIGGTCGCPVHKDRSCYICHROLLFCRVTLGKSFQFSA 1020
QY 1047 MKMAHSPGHHSTGRPSVNGLAALAEYVYIRGEQAYPEYLITYQIMRPEGMVDG 1100
Db 1021 MKMAHSPGHHSTGRPSVNGLAALAEYVYIRGEQAYPEYLITYQIMRPEGMVDG 1074

RESULT 11
AAB47022
ID AAB47022 standard; Protein; 1166 AA.
XX
AC AAB47022;
XX
DT 29-MAR-2001 (first entry)
XX
DE Human SPANK.

```



XX SPANK: SAM: sterile alpha motif; PARG: insulin resistance;  
 KW poly adenosine diphosphate-ribose polymerase; catalytic domain;  
 KW ANK: ankyrin repeat; cytosol; insulin-responsive aminopeptidase;  
 KW IRAP: GLUT4; adipocyte; insulin signalling pathway; hyperlipidaemia;  
 KW glucose intolerance; atheromatous disease; atherosclerosis;  
 KW obesity; cardiac insufficiency; coronary insufficiency; stroke;  
 KW high blood pressure; non-insulin dependent diabetes; hypertension;  
 KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 173..209  
 FT /note= "Encoded by nucleotides 768..878"  
 FT Misc-difference 327..362  
 FT /note= "Encoded by nucleotides 1230..1337"  
 FT Misc-difference 375..398  
 FT /note= "Encoded by nucleotides 1374..1445"  
 FT Misc-difference 482..524  
 FT /note= "Encoded by nucleotides 1695..1823"  
 FT Misc-difference 641..677  
 FT /note= "Encoded by nucleotides 2172..2282"  
 PN W0200077225-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 09-JUN-2000; 2000WO-US15926.  
 XX  
 PR 11-JUN-1999; 99US-0138957.  
 XX  
 PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Chi N, Lodish HF;  
 XX  
 XX WPI: 2001-091404/10.  
 DR N-PSDB; AAC85294.  
 XX  
 PT New insulin signalling protein SPANK, useful for reducing body mass,  
 PT glucose intolerance or insulin resistance and for preventing or  
 PT treating obesity-related and muscle-related diseases  
 XX  
 PS Claim 3; Fig 3; 65pp; English.  
 XX  
 CC This sequence represents human SPANK. The SPANK protein comprises  
 CC 3 domains:  
 CC (a) a SAM (sterile alpha motif) domain;  
 CC (b) a PARG (poly adenosine diphosphate-ribose polymerase) catalytic  
 CC domain; and  
 CC (c) an ANK domain composed of ankyrin repeats.  
 CC SPANK is a cytosolic protein which can poly(ADP-ribose)late itself.  
 CC SPANK binds insulin-responsive aminopeptidase (IRAP) and modulates  
 CC translocation of GLUT4 in the perinuclear region of adipocytes. It  
 CC is an effector in the insulin signalling pathway in eukaryotic cells.  
 CC SPANK is useful for reducing body mass, reducing glucose  
 CC intolerance or insulin resistance, for preventing or treating  
 CC obesity-related diseases or disorders, such as obesity, cardiac  
 CC insufficiency, coronary insufficiency, stroke, hypertension,  
 CC atheromatous disease, atherosclerosis, high blood pressure, non-insulin  
 CC dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is  
 CC also useful for preventing or treating muscle-related diseases or  
 CC disorders, such as muscular dystrophy, muscle atrophy and muscle  
 CC fatigue. Antibodies immunospecific for SPANK are useful for detecting  
 CC the presence of SPANK polypeptide in a biological sample.  
 XX  
 SQ Sequence 1166 AA;

Query Match 90.2%; Score 5201; DB 22; Length 1166;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1025; Conservative 1; Mismatches 0; Indels 148; Gaps 10;

QY 1 GFGKDVVEYLQNGASVQARDGGILPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60  
 DB 67 GFGKDVVEYLQNGASVQARDGGILPLHNACSFHAEVNVNLLRHGADPNARDNNYT 126  
 QY 61 PLHEAAIKGKIDVICIVLLQHGAEPTIRNTOGRTALDLPDSAKAVL----- 106  
 DB 127 PLHEAAIKGKIDVICIVLLQHGAEPTIRNTOGRTALDLPDSAKAVLKMALLPLVNWCH 186  
 QY 107 ---TCEYKDELLESARSNEEKMMALLPLVNVCHASDGRKSTPLHLAAGYNRVYIQ 162  
 DB 187 ASDGTGETYKDELLESARSNEE-----RKSTPLHLAAGYNRVYIQ 228  
 QY 163 LLLQHGADVHAKDKGDLPLHNACSYGHEVTELLVKGACVNMADLMQFTPLHEAASKN 222  
 DB 229 LLLQHGADVHAKDKGDLPLHNACSYGHEVTELLVKGACVNMADLMQFTPLHEAASKN 288  
 QY 223 RVVCSLLSYGADPTLLCHNKSALDAPTPOLKERL-----AYEFKG 266  
 DB 289 RVVCSLLSYGADPTLLCHNKSALDAPTPOLKERL-----AYEFKG 348  
 QY 267 HSLQQAAREADVTRIKKHLSEMVNFNFKPQTHETALHCAAAAPYPRKKOICEYLLRKGAN 326  
 DB 349 HSLQQAAREADVTR-----TETALHCAAAAS---RKQICEYLLRKGAN 389  
 QY 327 INEKT-----EFLTPLVHASEKAHNDVVEVVKHEAKVNALDNLGOTSLHRAAYCGHLOTC 383  
 DB 390 INEKT-----EFLTPLVHASEKAHNDVVEVVKHEAKVNALDNLGOTSLHRAAYCGHLOTC 449  
 QY 384 RLLLSYCGDPNLSLOGFTALQNGENNVQOLL-----QEGISLGNSE 425  
 DB 450 RLLLSYCGDPNLSLOGFTALQNGENNVQOLL-----QEGISLGNSE 509  
 QY 426 ADROLLEAAKAGDVETVKKLTQVSVNCRDIEGROSTPLHFAAGYNRVSVYELLQHGAD 485  
 DB 510 ADROLLEAAKAGDVE-----ROSTPLHFAAGYNRVSVYELLQHGAD 551  
 QY 486 VHAOKGGLVPLHNACSYGHEVAELLVKGAVNVADLMKFTPLHEAAAKGYEICKLL 545  
 DB 552 VHAOKGGLVPLHNACSYGHEVAELLVKGAVNVADLMKFTPLHEAAAKGYEICKLL 611  
 QY 546 LQHGADPTKNRDGNTPDLVDKGDTDIQLDLRGDAALLDAAKKGCLARVKKLSSPDVNV 605  
 DB 612 LQHGADPTKNRDGNTPDLVDKGDTDIQLDLRGDAALLDAAKKGCLARVKKLSSPDVNV 652  
 QY 606 CRDTQ-----RHSSTPLHAAAGYNNLEVAEYLLQHGADVNAODKGG 646  
 DB 653 CRDTQ-----RHSSTPLHAAAGYNNLEVAEYLLQHGADVNAODKGG 712  
 QY 647 LIPLHNAASYGHVDVAALLIKYNACVNTDKWAFPLHEAAQGRFTQCALLLAHGADPT 706  
 DB 713 LIPLHNAASYGHVDVAALLIKYNACVNTDKWAFPLHEAAQGRFTQCALLLAHGADPT 772  
 QY 707 LKQNGQTPDLVDVADVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSP 766  
 DB 773 LKQNGQTPDLVDVADVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSP 832  
 QY 767 SSLSAASLDNLSSGFSSELSSVSSSGTEGASLEKKEVPCVDFTSITQFVRNLGLEHMD 826  
 DB 833 SSLSAASLDNLSSGFSSELSSVSSSGTEGASLEKKEVPCVDFTSITQFVRNLGLEHMD 892  
 QY 827 IPEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPYTLNTSSG 886  
 DB 893 IPEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPYTLNTSSG 952  
 QY 887 TILDLSPDDKQFQSVSEEMQSTVRHRRDGGHAGIENRYNLTOKVCNKKLWERYTHR 946  
 DB 953 TILDLSPDDKQFQSVSEEMQSTVRHRRDGGHAGIENRYNLTOKVCNKKLWERYTHR 1012  
 QY 947 RKEYSENNHNAERMLFHGSPFVNAIHKGFDRHAYIGMFGAGIYFAENSKSNQYV 1006  
 DB 1013 RKEYSENNHNAERMLFHGSPFVNAIHKGFDRHAYIGMFGAGIYFAENSKSNQYV 1072  
 QY 1007 YGIGGGTGPVHKHRSYCHICRHLRFLCRVTVLGSFLQFSAMKMAHSPGHHVTVGRPSVN 1066

27-1  
 setch for  
 10/26/10  
 10/26/10

```

Db 1073 YGIGGGTGVHDKRSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPPGHSHVTVGRPSVN 1132
QY 1067 GLAIAEYVIYRGQAYPEYLITYQIMRPGWVDG 1100
Db 1133 GLAIAEYVIYRGQAYPEYLITYQIMRPGWVDG 1166

RESULT 12
AAB27209
ID AAB27209 standard; Protein: 1333 AA.
AC AAB27209;
XX
XX 27-FEB-2001 (first entry)
XX
XX Human tankyrase II SEQ ID NO: 2.
XX Human; tankyrase II; telomere length; signal transduction.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Domain 1..180
FT /note= "ANK domain"
FT Domain 181..1017
FT /note= "ANK domain"
FT Misc-difference 199
FT /label= Xaa
FT /note= "encoded by CNN"
FT Misc-difference 208..209
FT /label= XaaXaa
FT /note= "encoded by ANCCNC"
FT Misc-difference 212
FT /label= Xaa
FT /note= "encoded by GNG"
FT Misc-difference 232
FT /label= Xaa
FT /note= "encoded by NTT"
FT Misc-difference 242
FT /label= Xaa
FT /note= "encoded by NTC"
FT Misc-difference 244
FT /label= Xaa
FT /note= "encoded by NCA"
FT Misc-difference 255
FT /label= Xaa
FT /note= "encoded by ANT"
FT Misc-difference 264
FT /label= Xaa
FT /note= "encoded by ANT"
FT Misc-difference 726
FT /label= Xaa
FT /note= "encoded by TNT"
FT Misc-difference 794
FT /label= Xaa
FT /note= "encoded by GAN"
FT Misc-difference 800
FT /label= Xaa
FT /note= "encoded by NTG"
FT Misc-difference 822..823
FT /label= XaaXaa
FT /note= "encoded by TTTNTNT"
FT Domain 1030..1095
FT /note= "SAM domain"
FT Misc-difference 1055
FT /label= Xaa
FT /note= "encoded by NAG"
FT /note= "Xaa-unknown"
FT Domain 1190..1328
FT /note= "PARP domain"
XX
XX WO200061813-A1.
PN

```

```

XX 19-OCT-2000.
XX
XX 10-APR-2000; 2000WO-US09558.
XX
XX 09-APR-1999; 99US-0128577.
XX 13-APR-1999; 99US-0129123.
XX (GERO-) GERON CORP.
XX
XX Morin GB, Funk WD, Piatyszek MA;
XX WPI; 2000-679503/66.
XX N-PSDB; AAC66823.
XX
XX Novel mammalian Tankyrase II polypeptide and the polynucleotide
XX encoding the polypeptide useful for modulating or maintaining telomere
XX length, replicative capacity, apoptosis, chromosome packing or gene
XX expression
XX
XX Disclosure; Fig 2; 52pp; English.
XX
XX The present sequence is a version of the human tankyrase II protein
XX sequence. The protein is thought to be involved in signal transduction in
XX the cell, and to have binding activity for other telomere-associated
XX proteins. It is possible that it plays a role in the regulation of
XX telomere length, thus affecting the replicative ability of the cell. The
XX protein is useful for ribosylating target proteins, for determining
XX tankyrase II binding activity in a sample, and for modulating telomere
XX length in a cell.
XX
XX Sequence 1333 AA;
XX
Query Match 85.5%; Score 4932.5; DB 21; Length 1333;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 963; Conservative 8; Mismatches 129; Indels 9; Gaps 2;
QY 1 GFGRKDVVEYLLQNGASVQARDGGLIPLHNACSFHGAEVYVLLRHGADPNARDNNWT 60
Db 225 GFGRKLNXYLLQNGANKXQLYNGGLIPLHNACSFHGAEVYVLLRHGADPNARDNNWT 284
QY 61 PLHEAALKGKIDVCIVLLQHGAEPTIRNTDGTALDLDPSAKAVLTGEYKKDELLESAR 120
Db 285 PLHEAALKGKIDVCIVLLQHGAEPTIRNTDGTALDLDPSAKAVLTGEYKKDELLESAR 344
QY 121 SGNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNVVKIVOLLQHGADVHAKDKGLV 180
Db 345 SGNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNVVKIVOLLQHGADVHAKDKGLV 404
QY 181 PLHNACSYGHEVTELLVKGACVNAIMDLWQFTPLHEAASKNRVVCSSLISYGADPTLL 240
Db 405 PLHNACSYGHEVTELLVKGACVNAIMDLWQFTPLHEAASKNRVVCSSLISYGADPTLL 464
QY 241 NCHNKSALDAPTOLKERLAYEFKHSLLQAAREADVTRIKKHLSEMVNFKHPQTHET 300
Db 465 NCHNKSALDAPTOLKERLAYEFKHSLLQAAREADVTRIKKHLSEMVNFKHPQTHET 524
QY 301 ALHCAASPYPKRKQICELLRLRGANINEKTEPLHVASEKAHNDVVEVVVKHEAV 360
Db 525 ALHCAASPYPKRKQICELLRLRGANINEKTEPLHVASEKAHNDVVEVVVKHEAV 584
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNILISLOGFTALQMGNNVQQLQEGIS 420
Db 585 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNILISLOGFTALQMGNNVQQLQEGIS 644
QY 421 LGNSEADROLLEAAKAGDVETVKKLCVQSVNCEDETEGROSTPLHFAAGNVRVVEYLL 480
Db 645 LGNSEADROLLEAAKAGDVETVKKLCVQSVNCEDETEGROSTPLHFAAGNVRVVEYLL 704
QY 481 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAIMDLWQFTPLHEAASKNRV 540
Db 705 QHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAIMDLWQFTPLHEAASKNRV 764

```

Qy 541 ICKLLQHGADPTKKNRDNTPDLVDKGDPTDIODLLRGDAALLDAAKGCLARVKKLSS 600  
Db 765 ICKLLQHGADPTKKNRDNTPDLVDKGDPTDIODLLRGDAALLDAAKGCLARVKKXXF 824  
Qy 601 PDVNCRDTOGRHSTPLHLA--AGYNNLEVAEYLLQHGADVNAODKGLPLHNAASYGH 658  
Db 825 PDVNCRDTOGRHSTPLHLACGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 884  
Qy 659 VDVAALLIKYNACVNATDKWFTPLHFAAQGRFQLCALLAHAGADPTLKNQEQOTPLDL 718  
Db 885 XX 944  
Qy 719 VSADVSALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSLDNL 778  
Db 945 XXXXXXXXXXXTAAMPSPVLPSCNKPQVLNGVRSFGATADALSSGSPSSLSAASSLDNL 1004  
Qy 779 SGFSFELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHMDIFEREQITLDVL 838  
Db 1005 SGFSFELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHMDIFXREQITLDVL 1064  
Qy 839 VEMGHKELKEIGINAYGHRHKLIGVERLISGQOGLNPYLTNTSGSGTILIDLSDDKE 898  
Db 1065 VEMGHKELKEIGINAYGHRHKLIGVERLISGQOGLNPYLTNTSGSGTILIDLSDDKE 1124  
Qy 899 FQSVSEEMQSTVREHRDGHAGGIFNRYNLIKIOKVCNK-----KLWERYTHRRKEVS 951  
Db 1125 FQSVSEEMQSTVREHRDGHAGGIFNRYNLIKIOKVCNKXXXXXXAKIRHEERYTHRRKEVS 1184  
Qy 952 EBNHNHNERMLFHGSPFVNATIIHKGFDERHAYTGGMFGAGIYFAENSCKSNQVYVGTG 1011  
Db 1185 EBNHNHNERMLFHGSPFVNATIIHKGFDERHAYTGGMFGAGIYFAENSCKSNQVYVGTG 1244  
Qy 1012 GTGCPVHKDRSCYICHROLLFCRVTLTKSFLQFSAMKMAHSPGPHSHSVTGPRSVNGLALA 1071  
Db 1245 GTGCPVHKDRSCYICHROLLFCRVTLTKSFLQFSAMKMAHSPGPHSHSVTGPRSVNGLALA 1304  
Qy 1072 EVVIYRGEQAYPEYLITYQIMRPGMVDG 1100  
Db 1305 EVVIYRGEQAYPEYLITYQIMRPGMVDG 1333  
RESULT 13  
AAB27212  
XX AAB27212 standard; Protein; 1327 AA.  
AC AAB27212;  
XX  
DT 27-FEB-2001 (first entry)  
XX  
DE Human tankyrase I protein sequence SEQ ID NO: 8.  
XX  
KW Human; tankyrase II; telomere length; signal transduction.  
XX  
OS Homo sapiens.  
XX  
PN WO200061813-A1.  
XX  
PD 19-OCT-2000.  
XX  
PF 10-APR-2000; 2000WO-US09558.  
XX  
PR 09-APR-1999; 99US-0128577.  
XX  
PR 13-APR-1999; 99US-0129123.  
XX  
PA (GERO-) GERON CORP.  
XX  
PI Morin GB, Funk WD, Piatyszek MA;  
XX  
DR WPI; 2000-679503/66.  
XX  
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide  
PT encoding the polypeptide useful for modulating or maintaining telomere  
PT length, replicative capacity, apoptosis, chromosome packing or gene

expression -  
Claim 4; Fig 5; 52pp; English.  
The present sequence is a version of the human tankyrase I protein  
sequence. The invention relates to the isolation of the protein and  
coding sequences of human tankyrase II. This protein is thought to be  
involved in signal transduction in the cell, and to have binding activity  
for other telomere-associated proteins. It is possible that it plays a  
role in the regulation of telomere length, thus affecting the replicative  
ability of the cell. The protein is useful for ribosylating target  
proteins, for determining tankyrase II binding activity in a sample, and  
for modulating telomere length in a cell.  
SQ Sequence 1327 AA;  
Query Match 84.8%; Score 4889.5; DB 21; Length 1327;  
Best Local Similarity 83.0%; Pred. No. 0;  
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;  
Qy 1 GFRKRDVVEYLLQNGASVOARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60  
Db 225 GFRKRDVVEYLLQNGASVOARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 284  
Qy 61 PLHEAATKGIKIDYCVILVQHGAEPTIRNTDGTALDAPSAKAVLTGEYKDKDELLESAR 120  
Db 285 PLHEAATKGIKIDYCVILVQHGAEPTIRNTDGTALDAPSAKAVLTGEYKDKDELLESAR 344  
Qy 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRKIVQLLQHGADVHAKDGLV 180  
Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRKIVQLLQHGADVHAKDGLV 404  
Qy 181 PLHNACSYGHEYTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240  
Db 405 PLHNACSYGHEYTELLVKGACVNMADLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 464  
Qy 241 NCHNKSADLAPTOLKRLAYEFKGHSLLOAAAREADVTRIKKHLSELMVNFKHPOHET 300  
Db 465 NCHGKSADVADNAPTPELRLTYEFKGHSLLOAAAREADVTRIKKHLSELMVNFKHPOHET 524  
Qy 301 ALHCAASAPYKPKQICELLRLKGANINEKTKEFTPLHVASEKAHNDVVEVVKHEAV 360  
Db 525 ALHCAVASLHPKPKQVTELLRLKGANVNEKNKDFMTPLHVAERAHNDVVEVVKHGAHM 584  
Qy 361 NALDNLGOTSILHRAAYCGHLQTCRLLSYCGDPNIIISLOGFTALQNGENNVQILQEGIS 420  
Db 585 NALDTLQOTALHRAALAGHLQTCRLLSYSGDPSIIISLOGFTAAQMGNEAVQILSESTP 644  
Qy 421 LGNSEADRLLEAAKAGDVEVTKLCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480  
Db 645 IRTSDVDYRLLEASKAGDLETVCQSSQNVNCRDLRGHSTPLHFAAGYNRVSVVEYLL 704  
Qy 481 QHGADVHAKDGLVPLHNACSYGHEVAELLYKHGAVNVNADLWKFPTLHEAASKGYE 540  
Db 705 HHGADVHAKDGLVPLHNACSYGHEVAELLYKHGAVNVNADLWKFPTLHEAASKGYE 764  
Qy 541 ICKLLQHGADPTKKNRDNTPDLVDKGDPTDIODLLRGDAALLDAAKGCLARVKKLSS 600  
Db 765 ICKLLQHGADPTKKNRDNTPDLVDKGDPTDIODLLRGDAALLDAAKGCLARVKKLCT 824  
Qy 601 PDVNCRDTOGRHSTPLHLAGYNNLEVAEYLLQHGADVNAODKGLPLHNAASYGHVD 660  
Db 825 PENINCRDQTGRNSTPLHLAGYNNLEVAEYLLPHGADVNAODKGGIPLHNAASYGHVD 884  
Qy 661 VAALLIKYNACVNATDKWFTPLHFAAQGRFQLCALLAHAGADPTLKNQEQOTPLDLS 720  
Db 885 IAALLIKYNTCVNATDKWFTPLHFAAQGRFQLCALLAHAGADPTLKNQEQOTPLDLS 944  
Qy 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASSLDNL 780  
Db 945 ADDIRALLIDAMPPEALPTCFKPOAT-----VVSASLISPASTPSCLSAASSIDNL 997  
Qy 781 SFSELSSVSSSGTEGASSLEKK--EVPGVDFSTQFVRNLGLEHMDIFEREQITLDVL 838

```

Db 998 PLAEAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIETETEITLDVL 1057
QY 839 VEMGHKELKEGINAYGHRHKLKGVRLISGQGLNPYLTLNTSGSGTILIDLSPPDKE 898
Db 1058 ADMGHEELKEGINAYGHRHKLKGVRLISGQGLNPYLTLNTSGSGTILIDLSPPDKE 1117
QY 899 FQSVVEEMQSTVREHRDGGHAGGIFNRYNLIKQKVGKKNKLWERYTHRRKEVSEENHHA 958
Db 1118 YQSVVEEMQSTVREHRDGGHAGGIFNRYNRIQKVNKKLRERFCHQKEVSEENHHA 1177
QY 959 NERMLFHGSPFVNAIHHKGFDERHAYTGGMFGAGIYFAENSSKSNQVYIGGTCGCPVH 1018
Db 1178 NERMLFHGSPFVNAIHHKGFDERHAYTGGMFGAGIYFAENSSKSNQVYIGGTCGCPH 1237
QY 1019 KDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGCHSHVGTGRPSVNGLALAEYVIYRG 1078
Db 1238 KDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGCHSHVGTGRPSVNGLALAEYVIYRG 1297
QY 1079 EOAYPEYLITYQIMRPE 1095
Db 1298 EOAYPEYLITYQIMKPE 1314

```

## RESULT 14

AAI44402  
ID AAY44402 standard; protein; 1327 AA.

AC AAY44402;

DT 22-MAR-2000 (first entry)

XX Human tankyrase.

DE Human tankyrase; TRF1, telomeric repeat binding factor-1;  
KW two-hybrid screen; telomere length regulation; recombinant tankyrase;  
KW aging; skin atrophy; macular degeneration; atherosclerosis;  
KW ataxia telangiectasia; tankyrase modulator; rational drug design.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 181..1010  
FT Domain /note= "Ankyrin-specific (ANK) repeat consensus domain"  
FT 1023..1088  
FT Domain /label= Sterile\_alpha\_motif  
FT 1176..1314  
FT Domain /note= "Poly(ADP-ribose)polymerase (PARP) related domain"

XX WO9964606-A1.

XX 16-DEC-1999.

XX 09-JUN-1999; 99WO-US12968.

XX 10-JUN-1998; 98US-0095225.

XX 17-AUG-1998; 98US-0135233.

XX 19-NOV-1998; 98US-0196387.

XX (UYRQ ) UNIV ROCKEFELLER.

XX De Lange T, Smith S;

XX WPI; 2000-116549/10.

DR N-PSDB; AAZ29627.

PT New nucleic acid encoding vertebrate tankyrase, a regulator of telomere  
PT length, used to identify modulators, e.g. for inhibiting growth of  
PT cancer  
XX Claim 1; Fig 1B; 133pp; English.

XX The present sequence is human tankyrase, a TRF1 (telomeric repeat binding  
XX

CC factor-1) binding protein. This was isolated using a two-hybrid screen  
CC with TRF1. Tankyrase is involved in regulation of telomere length through  
CC TRF1 and may directly modulate the effect of TRF1. Recombinant tankyrase,  
CC or its fragments are used to identify specific modulators which are  
CC potential drugs for counteracting telomere shortening associated with aging  
CC (e.g. atrophy of skin, macular degeneration or atherosclerosis) or  
CC diseases like ataxia telangiectasia. They are also used to raise specific  
CC antibodies (used as immunoassay reagents and as modulators of tankyrase  
CC activity) and in rational drug design.

XX Sequence 1327 AA;

Query Match 84.8%; Score 4889.5; DB 21; Length 1327;  
Best Local Similarity 83.0%; Pred. No. 0;  
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

```

QY 1 GFKRKDVVEYLLONGASVOARDGGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYIT 60
Db 225 GFKRKDVVEYLLONGASVOARDGGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYIT 284
QY 61 PLHEAAIKGKIDVCIVLLOHGADEPTINTDGRALDADPSAKAVLTGEYKKDELLESAR 120
Db 285 PLHEAAIKGKIDVCIVLLOHGADEPTINTDGRALDADPSAKAVLTGEYKKDELLESAR 344
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 180
Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGDLV 404
QY 181 PLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEYVCSLLSYGADPTLL 240
Db 405 PLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRVEYVCSLLSYGADPTLV 464
QY 241 NCHNKSATDLCAPTPOLKERLAYEFKSHLSLOAAREADVTRIKHLSLEMYNFKHPQTHET 300
Db 465 NCHNKSATDLCAPTPOLKERLAYEFKSHLSLOAAREADVTRIKHLSLEMYNFKHPQTHET 524
QY 301 ALHCAASPPYKPKQICELLLRKGANINEKTEFLPLHVASEKAHNDVYVYVVKHEAKV 360
Db 525 ALHCAASPPYKPKQICELLLRKGANINEKTEFLPLHVASEKAHNDVYVYVVKHEAKV 584
QY 361 NALDNLGQTSIHLRAAYCGHLOTCRLLSYGCDDNITISLQGTALOMGNENVOQLQSGIS 420
Db 585 NALDNLGQTSIHLRAAYCGHLOTCRLLSYGCDDNITISLQGTALOMGNENVOQLQSGIS 644
QY 421 LGSSEADROLLEAAKAGDVETVKLCTVQSVNCRDLEGROSTPLHFAAGYVRYVVEYLL 480
Db 645 LGSSEADROLLEAAKAGDVETVKLCTVQSVNCRDLEGROSTPLHFAAGYVRYVVEYLL 704
QY 481 QHGADVHAKDKGGLVPLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRYE 540
Db 705 QHGADVHAKDKGGLVPLHNACSYGHYEYVELLVKHGACVNMADLWQFTPLHEAASKNRYE 764
QY 541 ICKLLLOHGADEPTKKNRDGNTPLDLVKDGDFTDQDLRGDAALLDAKKGCLARVKKLSS 600
Db 765 ICKLLLOHGADEPTKKNRDGNTPLDLVKDGDFTDQDLRGDAALLDAKKGCLARVKKLSS 824
QY 601 PDVNCNCRDQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
Db 825 PDVNCNCRDQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 884
QY 661 VAALLIKYNACVNAATDKWAFTPLHEAAQKRTQALCALLAHGADEPTLKNOEGOTPLDLVS 720
Db 885 VAALLIKYNACVNAATDKWAFTPLHEAAQKRTQALCALLAHGADEPTLKNOEGOTPLDLVS 944
QY 721 ADDVSALLTAAMPSPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLNLG 780
Db 945 ADDVSALLTAAMPSPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLNLG 997
QY 781 SFSESSVSSVSSGTEGASSLEKK--EVPGVDFSTITQVFNGLGLEHMDIFERQITLDVL 838
Db 998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIETETEITLDVL 1057
QY 839 VEMGHKELKEGINAYGHRHKLKGVRLISGQGLNPYLTLNTSGSGTILIDLSPPDKE 898

```



Wed Feb 12 16:08:51 2003

us-09-843-159b-3.rag

Page 16

Search completed: February 12, 2003, 03:32:18  
Job time : 43.6068 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 07:42:50 : Search time 15,0427 Seconds  
(without alignments)  
2151.551 Million cell updates/sec

Title: US-09-843-159B-3  
Perfect score: 5769  
Sequence: 1 GFGRKDVVEYLLONGASVQA.....AYPEYLITYQIMRPGMVDG 1100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCFUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5717	99.1	1166	4	US-09-350-982C-5
2	4889.5	84.8	1327	4	US-09-196-387-2
3	3093	53.6	949	4	US-09-196-387-10
4	1918.5	33.3	673	4	US-09-196-387-8
5	860.5	14.9	1839	2	US-09-172-977-4
6	841.5	14.6	843	2	US-09-172-977-3
7	821	14.2	1745	2	US-09-031-485-33
8	821	14.2	1745	2	US-08-847-429A-33
9	821	14.2	1745	3	US-09-065-474-33
10	821	14.2	1745	4	US-09-557-034-33
11	629.5	10.9	1088	4	US-09-082-059-2
12	513	8.9	352	3	US-09-065-474-139
13	513	8.9	352	4	US-09-557-034-139
14	421.5	7.3	303	2	US-09-031-485-23
15	421.5	7.3	303	2	US-08-847-429A-23
16	421.5	7.3	303	3	US-09-065-474-23
17	421.5	7.3	303	4	US-09-557-034-23
18	409	7.1	302	2	US-09-031-485-38
19	409	7.1	302	2	US-08-847-429A-38
20	409	7.1	302	3	US-09-065-474-38
21	409	7.1	302	4	US-09-557-034-38
22	392	6.8	348	2	US-09-031-485-28
23	392	6.8	348	2	US-08-847-429A-28
24	392	6.8	348	3	US-09-065-474-28
25	392	6.8	348	4	US-09-557-034-28
26	390.5	6.8	787	4	US-09-188-930-334
27	379.5	6.6	1719	4	US-09-605-785-378

ALIGNMENTS

RESULT 1

US-09-350-982C-5  
: Sequence 5, Application US/09350982C

: Patent No. 6455290

: GENERAL INFORMATION:

: APPLICANT: Bertheisen, Jens

: APPLICANT: Toma, Salvatore

: TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R

: TITLE OF INVENTION: Same

: FILE REFERENCE: PHRM-0043

: CURRENT APPLICATION NUMBER: US/09/350,982C

: CURRENT FILING DATE: 1999-07-09

: NUMBER OF SEQ ID NOS: 10

: SOFTWARE: PatentIn version 3.0

: SEQ ID NO 5

: LENGTH: 1166

: TYPE: PRT

: ORGANISM: Artificial

: FEATURE:

: NAME/KEY: misc\_feature

: OTHER INFORMATION: Xaa is any amino acid

: NAME/KEY: misc\_feature

: LOCATION: (1102)..(1102)

: OTHER INFORMATION: n is any nucleic acid

: NAME/KEY: misc\_feature

: LOCATION: (2650)..(2650)

: OTHER INFORMATION: n is any nucleic acid

: US-09-350-982C-5

Query Match 99.1%; Score 5717; DB 4; Length 1166;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1091; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy	1	GFGRKDVVEYLLONGASVQA	RDGGLIPLHNACSF	GHAEEVYNLLRHGADPNARDNNYT	60
Db	67	GFGRKDVVEYLLONGANVQ	ARDGGLIPLHNACSF	GHAEEVYNLLRHGADPNARDNNYT	126
Qy	61	PLHEAAIKGKIDVCI	VLILQHGAEPTIR	NTDGTALDLP	120
Db	127	PLHEAAIKGKIDVCI	VLILQHGAEPTIR	NTDGTALDLP	186
Qy	121	SGNEKKMALLTPLNVN	CHASDGRKSTPL	HLAAGYNRVKIV	180
Db	187	SGNEKKMALLTPLNVN	CHASDGRKSTPL	HLAAGYNRVKIV	246
Qy	181	PLHNACSYGHVEVTE	LLVKHGACYNAMD	WQFTPLHEAASKNR	240
Db	247	PLHNACSYGHVEVTE	LLVKHGACYNAMD	WQFTPLHEAASKNR	306

QY 241 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 300  
Db NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 366  
QY 301 ALHCAASPYPKRQICELLRRKGANINERTKFLTPLVHASEKAHNDVVEVVVKHEAKV 360  
Db AXHCAASPYPKRQICELLRRKGANINERTKFLTPLVHASEKAHNDVVEVVVKHEAKV 426  
QY 361 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDDNITISLOGFTALOMGNENVOQLLOEGIS 420  
Db NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDDNITISLOGFTALOMGNENVOQLLOEGIS 486  
QY 421 LGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480  
Db LGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 546  
QY 481 OHGADVHAKDKGLVPLHNAACSYGHEVEABLLVKGAVNVADLWFTPLHFAAAGKYE 540  
Db OHGADVHAKDKGLVPLHNAACSYGHEVEABLLVKGAVNVADLWFTPLHFAAAGKYE 606  
QY 541 ICKLLQHGADPTKKNRDGNTPLDLVKDGTDTIDQLLRGDAALLDAKGLARVKKLSS 600  
Db ICKLLQHGADPTKKNRDGNTPLDLVKDGTDTIDQLLRGDAALLDAKGLARVKKLSS 666  
QY 601 PDNVNCRDTQGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 660  
Db PDNVNCRDTQGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 726  
QY 661 VAALLIKYNACVNATDKWFTPLHFAAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 720  
Db VAALLIKYNACVNATDKWFTPLHFAAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYCHVD 786  
QY 721 ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLG 780  
Db ADDVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLG 846  
QY 781 SFSELSVSSVSSGTEGASLEKEVPDVFSTQFVNRNLGLEHMDIFEREQITLDVLVE 840  
Db SFSELSVSSVSSGTEGASLEKEVPDVFSTQFVNRNLGLEHMDIFEREQITLDVLVE 906  
QY 841 MGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPLTLNTSGSTILIDLSPPDKFEQ 900  
Db MGHKELKEIGINAYGHRHKLKIGVERLISGQGLNPLTLNTSGSTILIDLSPPDKFEQ 966  
QY 901 SVEEEMQSTVREHRDGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNHANE 960  
Db SVEEEMQSTVREHRDGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNHANE 1026  
QY 961 RMLFHGSPFNALIHKGDPDERHAYIGGMPGAGIYFAENSCKSNQYVYIGGTGCPVHKD 1020  
Db RMLFHGSPFNALIHKGDPDERHAYIGGMPGAGIYFAENSCKSNQYVYIGGTGCPVHKD 1086  
QY 1021 RSCYICHRQLLCFVRTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGLALAEVVIYRGQ 1080  
Db RSCYICHRQLLCFVRTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGLALAEVVIYRGQ 1146  
QY 1081 AYPEYLITYQIMRPGMVDG 1100  
Db AYPEYLITYQIMRPGMVDG 1166

## RESULT 2

US-09-196-387-2

; Sequence 2, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber &amp; Jackson

; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196,387  
; FILING DATE:  
; CLASSIFICATION:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/095,225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1327 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; US-09-196-387-2

Query Match 84.8%; Score 4889.5; DB 4; Length 1327;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

QY 1 GFRKDVVEYLLQNGASVQARDGGGLIPLHNACSFHAEVYNNLLLRHGADPNARDNNYT 60  
Db 225 GFRKDVVEYLLQNGASVQARDGGGLIPLHNACSFHAEVYNNLLLRHGADPNARDNNYT 284  
QY 61 PLHEAAIKKIDVCIVLLOHGAETIRNTDGTALDAPSAKAVLTGEVKDELLESAR 120  
Db 285 PLHEAAIKKIDVCIVLLOHGAETIRNTDGTALDAPSAKAVLTGEVKDELLESAR 344  
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRVRIQVLLQHGADVHAKDKGLV 180  
Db 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRVRIQVLLQHGADVHAKDKGLV 404  
QY 181 PLHNACSYGHEVTELLVYKHGACVNAQDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240  
Db 405 PLHNACSYGHEVTELLVYKHGACVNAQDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 464  
QY 241 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 300  
Db 465 NCHNKSALDAPTPOLKERLAYEFKHSLLQAAREADVTRIKKHLSELMVNFKHPOTHET 524  
QY 301 ALHCAASPYPKRQICELLRRKGANINERTKFLTPLVHASEKAHNDVVEVVVKHEAKV 360  
Db 525 ALHCAASPYPKRQICELLRRKGANINERTKFLTPLVHASEKAHNDVVEVVVKHEAKV 584  
QY 361 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDDNITISLOGFTALOMGNENVOQLLOEGIS 420  
Db 585 NALDNLGOTSLSHRAAYCGHLOTCLRLLSYGCDDNITISLOGFTALOMGNENVOQLLOEGIS 644  
QY 421 LGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 480  
Db 645 LGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLL 704  
QY 481 OHGADVHAKDKGLVPLHNAACSYGHEVEABLLVKGAVNVADLWFTPLHFAAAGKYE 540  
Db OHGADVHAKDKGLVPLHNAACSYGHEVEABLLVKGAVNVADLWFTPLHFAAAGKYE 540



Db 705 HHGADVHAKDGGGLVPLHNACSYGHEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764  
QY 541 ICKLLQHGADPTKKNRGNTPLDLVKDGDTDIDQLLRGDAALLDAKKGCLARVKKLSS 600  
Db 765 ICKLLKHGADPTKKNRGNTPLDLVKEGDTDIDQLLRGDAALLDAKKGCLARVQKLT 824  
QY 601 PDVNCRTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDGGGLPLHNAASYGHVD 660  
Db 825 PENINCRTQGRNSTPLHLAAGYNLEVAEYLLHAGADVNAQDGGGLPLHNAASYGHVD 884  
QY 661 VAALLIKYNACVNAATDKWAFPLHEAAQKGRQTCALLLAHAGDPTLKNQEGQTPDLVLS 720  
Db 885 IAALLIKYNTCVNATDKWAFPLHEAAQKGRQTCALLLAHAGDPTLKNQEGQTPDLVLS 944  
QY 721 ADDVALLTAAMPSPALSCYKQVNLNGVRSPCATADALSSGSPSSLSAASSLDNLGS 780  
Db 945 ADDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTPSCLSAASSIDNLG 997  
QY 781 SFSELSVVSSSGTSGASSLEKK--EVPQVDFSIQFVRNLGLEHMDIFEREQITLDVL 838  
Db 998 PLAELAVGASNAGDAAGTERKEGEVAGLDNINSQFLKSLGLEHLRDIFFETEQITLDVL 1057  
QY 839 VEMGHKELKEIGINAYHRHKLKIGVERLISQOGLNPYLTNTSGSGTILIDJSPDCKE 898  
Db 1058 ADMGHEELKEIGINAYHRHKLKIGVERLISQOGLNPYLTNPHFCVNOQTILIDJLAPEDKE 1117  
QY 899 FOSVEEMOSTVREHRDGGHAGGIFNRYNLIKQVCHKKLWERYTHRRKEVSEENHHA 958  
Db 1118 YOSVEEMOSTVREHRDGGHAGGIFNRYNIRIQVWAKKLRERCHROKEVSEENHNNH 1177  
QY 959 NERMFLPHGSPFNATIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGGTCPCPVH 1018  
Db 1178 NERMFLPHGSPFNATIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYGGGTCPCPTH 1237  
QY 1019 KDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGALAEVYVIRG 1078  
Db 1238 KDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSTVGRPSVNGALAEVYVIRG 1297  
QY 1079 EQAYPEYLITYQIMRPE 1095  
Db 1298 EQAYPEYLITYQIMRPE 1314

## RESULT 3

US-09-196-387-10

; Sequence 10, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber &amp; Jackson

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent In Release #1.0, Version #1.30

; APPLICATION NUMBER: US/09/196,387

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/095,225

; FILING DATE: June 10, 1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 949 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-196-387-10

Query Match 53.6%; Score 3093; DB 4; Length 949;

Best local Similarity 81.5%; Pred. No. 2e-272;

Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;

QY 1 GFGRKDVVEYLLQNGASVQARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60  
Db 225 GFGRKDVVEYLLQNGANVHARDGGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 284  
QY 61 PLHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLEAAR 120  
Db 285 PLHEAAIKGKIDVCIVLQHGADPNIRNTDGSALDADPSAKAVLTGEYKKDELLEAAR 344  
QY 121 SGNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNVKTIVOLLQHGADVHAKDKGDLV 180  
Db 345 SGNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNVKTIVOLLQHGADVHAKDKGDLV 404  
QY 181 PLHNACSYGVHEVTELVKKGACVNAAMDLMQFTPLHEAAAKNRVEVCSLLLSYGADPTLL 240  
Db 405 PLHNACSYGVHEVTELVKKGACVNAAMDLMQFTPLHEAAAKNRVEVCSLLLSYGADPTLV 464  
QY 241 NCHNKSAIDLAPTOLKRLAYEFKSHSLQAAREADVTRIKKLSLEMVNFKHPQTHET 300  
Db 465 NCHGKSAVDMAPTPELRERLTVEFKSHSLQAAREADLAKYKKTALAEIINFKQPOSHET 524  
QY 301 ALHCAAAAPYKPKOICELLRKGNANETKEFTPLHVASEKAHNDVVEVVKHEAVK 360  
Db 525 ALHCAVASLHPKPKOVTELLRKGNANETKEFTPLHVASEKAHNDVVEVVKHEAVK 584  
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLQGTALQMGNNVQQLQEGIS 420  
Db 585 NALDTLQGTALHRAALAGHLQTCRLLLSYGCSDPSIISLQGTAAQMGNEAVQQLSESTP 644  
QY 421 LGNSEADQLLEAAKAGDVETVKLCIVQSVNCRDIEGRQSTPLHFAAGYNRVSVWEYLL 480  
Db 645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDIEGRHSTPLHFAAGYNRVSVWEYLL 704  
QY 481 OHGADVHAKDKGGLVPLHNACSYGHEVAELLVKHGVNVNADLWKFTPLHEAAAKGKYE 540  
Db 705 HHGADVHAKDKGGLVPLHNACSYGHEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764  
QY 541 ICKLLQHGADPTKKNRGNTPLDLVKDGDTDIDQLLRGDAALLDAKKGCLARVKKLSS 600  
Db 765 ICKLLKHGADPTKKNRGNTPLDLVKEGDTDIDQLLRGDAALLDAKKGCLARVQKLT 824  
QY 601 PDVNCRTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDGGGLPLHNAASYGH-- 658  
Db 825 PENINCRTQGRNSTPLHLAAGYNLEVAEYLLHAGADVNAQDGGGLPLHNAASYGCL 884  
QY 659 VDVAALLIKYNACVNAATDKWAFPLHEAAQKGRQTCALLLAHAGDPTLKNQEGQTP 716  
Db 885 ARVQKLTCPENINCRTQGRNSTPLHLAAGYNLEVAEYLLHAGADVNAQDGGGLPL 942

## RESULT 4

US-09-196-387-8

; Sequence 8, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

APPLICANT: de Lange, Titia  
APPLICANT: Smith, Susan  
TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
TITLE OF INVENTION: OF USE THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,387  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,225  
FILING DATE: June 10, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 673 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-196-387-8

Query Match 33.3%; Score 1918.5; DB 4; Length 673;  
Best Local Similarity 80.9%; Pred. No. 1.1e-165;  
Matches 364; Conservative 42; Mismatches 43; Indels 1; Gaps 1;  
QY 1 GFGKDVVEYLLQNGASVQARDGGGLPLHNACSFSGHAEVYVNLRLRHGADPNARDNNYIT 60  
DB 225 GFGKDVVEYLLQNGASVQARDGGGLPLHNACSFSGHAEVYVNLRLRHGADPNARDNNYIT 284  
QY 61 PLHEAAIKGKIDVICVILLOHGAEPTIRNTDGRALDLPDSAKAVLTGEYKKDELLESA 120  
DB 285 PLHEAAIKGKIDVICVILLOHGAEPTIRNTDGRALDLPDSAKAVLTGEYKKDELLESA 344  
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 180  
DB 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLV 404  
QY 181 PLHNACSYGHEVTELLVKGACVNDLQWFTPLHEAASKNVEVCSLLLSYGADPTLL 240  
DB 405 PLHNACSYGHEVTELLVKGACVNDLQWFTPLHEAASKNVEVCSLLLSYGADPTLL 464  
QY 241 NCHNKAIDLAPTPOLKERLAYEFKGSLLQAAAREADVTRIKHLSLEWYVNFKHQTHET 300  
DB 465 NCHGKSAVDMAPTPELREUTYEFKGSLLQAAAREADVTRIKHLSLEWYVNFKHQTHET 524  
QY 301 ALHCAASPYPKRQKICELLRRKGANINEKTEFLTPHVAEKAHNDVVEVVKHEAVK 360  
DB 525 ALHCAVASLHPKQVTELLLRKGANVNEKNKDFMTPLHVAERAHNDVVEVVKHEAVK 584  
QY 361 NALDNLQVSLHRAAYCGHLQTCRLLLSYCGDPNIIISLOGFTALQMGNEVNVQQLQEGIS 420  
DB 585 NALDTLQGTALHRAALAGHLQTCRLLLSYCGDPNIIISLOGFTALQMGNEVNVQQLQEGIS 644  
QY 421 LGNSEADRLLEAAKAGDVETVKKLCTVQS 450

Db 645 SDPSIISLOGFTAAQMGNEVNVQQLQEGIS 673

RESULT 5  
US-09-172-977-4  
Sequence 4, Application US/09172977  
Patent No. 5989863  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
APPLICANT: Yue, Henry  
TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
FILE REFERENCE: PF-0615 US  
CURRENT APPLICATION NUMBER: US/09/172,977  
CURRENT FILING DATE: 1998-10-14  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PERL Program  
SEQ ID NO 4  
LENGTH: 1839  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: g29491  
US-09-172-977-4

Query Match 14.9%; Score 860.5; DB 2; Length 1839;  
Best Local Similarity 29.1%; Pred. No. 1.5e-68;  
Matches 297; Conservative 147; Mismatches 404; Indels 174; Gaps 29;  
QY 3 GRKDVVEYLLQNGASVQARDGGGLPLHNACSFSGHAEVYVNLRLRHGADPNARDNNYITPL 62  
DB 75 GHVGLVQELLGRGSSVDSATKKNTALHIASLAGOAEVYVNLREGANINAAQSQNGETPL 134  
QY 63 HEAAIKGKIDVICVILLOHGAEPTIRNTDGRALDLPDSAKAVLTGEYKKDEL--- 115  
DB 135 YMAAQENHIDVVKYLLENGANQSTATEDGGTPLAVALQOQHNOVAAILLENDTKGKRLP 194  
QY 116 -LESARSGNEEKKMMALLTPLNVNCHASDGRKS-----TPLHLAAGYNRVKIVQL 163  
DB 195 ALHIAARKDDTKSAALLQ---NDHNADVQSKMMVNRTEGFTPLHIAAHYGVNVVATL 251  
QY 164 LLOHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNDLQWFTPLHEAASKNR 223  
DB 252 LLNRAAVDTARNITPLHVASKRGNTVMVKLLDRGGQIDAKTRDGLTPLHCAARSGH 311  
QY 224 VECSLLLSYGADPTLLNCHNKAIDLAPTPOLKERLAYEFKGSLLQAAAREADVTRIK 283  
DB 312 DVVEVLELLERGA-PLIARTKN-----GLSPLHMAAQGDHVECVK 349  
QY 284 HLSLEWYVNFKHP-----QTHETALHCAASPYPKRQKICELLRRKGANINEKTEFLT 338  
DB 350 HL-----LQHKAPVDVDTLDYLTALHVA---HCGHYEVTKLLDKRANPARALNGETPL 402  
QY 339 HVASEKAHNDVVEVVKHEAKVNDLQWFTPLHRAAYCGHLQTCRLLLSYCGDPNIIISL 398  
DB 403 HIACKNNRIKVMELLYKYGASIQAITESGLTPIHVAAFMGHLNVLVLLONGASPDVTNI 462  
QY 399 QGFTALQMGNEVNVQQLQEGISLGNSEADRLLEAAKAGDVETVKKLCTVQSVNCRDIEG 458  
DB 463 RGETALHM-----AARAGOVVVR--CLLRNGALVDARA 494  
QY 459 R-OSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGA 517  
DB 495 REQOTPLHITASRLGKTEIVOLLQOHMAHPDAATNGTTPHLHISAREGQVDVASVLEAGA 554  
QY 518 VNVADLWFTPLHEAAKAGKYEICKLLQLQHGADPTKKNRDGNTPDLQVKGDTDTODLL 577  
DB 555 AHSLATKKGFTPLHVAAYKGSGLDVAKLLQRRAAADSAGKNGLTPLHVAHYDN----- 608  
QY 578 RGDALLDRAKKGCLARVKLSSPDVNVNCRDGTGRHSTPLHLAAGYNVNLVEAVYLLQHGA 637

Db 609 -----QKVALLLLEKASP-----HATAKNGYTPHLHIAAKKNQMOIASTLLNYGA 653  
QY 638 DYNADQKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTOCAL 697  
Db 654 ETNIVTKGVTPHLHSAOEGHTDMVTLLLDKGANIHMTSGSLTSLHAAQEDKVNADI 713  
QY 698 LLAHGADPTLKNQOGTDLIV-----SADDVSALLTAAMPSPA-LPSCYKP----- 743  
Db 714 LTKHGADODATKLGTYPLIVACHYGNVKNVFLKQGANVNAKTKNGYTPHLHQAQOQH 773  
QY 744 -OVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASLEK 802  
Db 774 THIIINVLLQHGAKPNATTA--NGNTALAIARLGYI--SVVDTLKVVVTEVTTTTTITE 829  
QY 803 KEVPGVDSITOFVRLGLEHLMDFERE---QITLDVLVEMGHKELKEIGINAYGHRHK 859  
Db 830 KHLNVPETMT-----VLVDSEEGDDTMTGGEYLRPEDLKELGDDS----- 874  
QY 860 LTKGVERLISGQ--QGLNPYLTNTSGSTILIDLSPDKKEFQSVSEEMOSTVREHRDGG 917  
Db 875 -----LPSSQFLDGMN-YLRYSLGEG-----RSDSLRSFSDRSHTLSHASYLRD-- 918  
QY 918 HAGGFNRYNLIKQVCN-KKLERYTHRRKEVSEENHNHANERMLFHGSPFFVNAIHK 976  
Db 919 --SAYMDDSVIPSHOVSTLAKAERNRYLSWGTENLDNVA-----LSSSP-----IHS 966  
QY 977 GF 978  
Db 967 GF 968

## RESULT 6

US-09-172-977-3

; Sequence 3, Application US/09172977  
; Patent No. 598863  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
; FILE REFERENCE: PF-0615 US  
; CURRENT APPLICATION NUMBER: US/09/172.977  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 843  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE: -  
; OTHER INFORMATION: g1841966  
US-09-172-977-3

Query Match 14.6%; Score 841.5; Db 2; Length 843;  
Best Local Similarity 29.7%; Pred. No. 2.1e-67;  
Matches 265; Conservative 138; Mismatches 351; Indels 137; Gaps 21;  
QY 3 GRKDVVEYLLONGASVQARDGGLIPLHNAASFGHAEVYNLLLRHGADPNARDNNYTP 62  
Db 40 GHVGLVQELLGSGVSDSATKGNNTALHIASLAGQAEVVKVILVKEGANINAOQSNGFTPL 99  
QY 63 HEAAIKGKIDCVILVQLQHGAEPTIRNTDGR-----ALDIADPSAKAVLTGEEKKDEL--- 115  
Db 100 YMAAQENHIDVVKVLLLENGANQSTATEDGFTPLAVALOQHQAVALLENDTKGKVRP 159  
QY 116 -LESASRGENEKKWALLTPLNVNCHASDGRKS-----TPHLAAGYNRVKIVOL 163  
Db 160 ALHIAARKDDTKSAALLQ---NDHNADVQSKMWNRTTESGFTPLHIAHYGNVNVATL 216  
QY 164 LLOHGADVHAKGDGLVPLHNAASYGHEVTELLVVKHACVNMADLWQFTPLHEAASKNR 223  
Db 217 LLNRGAADVFTARGITPLHVASKRGNTNMVKLLLDGRGGIDAKTRDGLTPLHCAARSGH 276

QY 224 VEYCSLLSYGADPTLLNCHNKSALDAPTQPKERLAYEFKGHSLLOAARADVYTRIKK 283  
Db 277 DQVVELLERGA--PLLARTKN-----GLSPHMAAQGDHVECVK 314  
QY 284 HUSLEMVNFKHP-----QTHETALHCAASPYPKRQICICELLKRCANINEXTKEPLTPL 338  
Db 315 HL-----LQHKAPVDDVTLDYLTALHVA--HCGHYRVTKLLLDKRNANRALNGFTPL 367  
QY 339 HVASEKAHNDVVEVYVYKHEAKVNALDNLGQTSLHRAAYGCHLQTCRLLSYSGDPNIIUL 398  
Db 368 HIACKNRKRYMBELLVKYGAIQAITESGLTPIPAAFMGHLNIVLLLLONGASPDVTNI 427  
QY 399 QGFTALOMGNENVOQLQEGISLGNSEADQRLLEAAKAGDVETVKKCTVQSVNCRDIEG 458  
Db 428 RGETALHM-----AARAGEVEVVR--CLLRNGALVDARA 459  
QY 459 R-QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNAASYGHEVAELLVKHGA 517  
Db 460 REEQTPLHIASRLGKTEIVOLLQHMAHPDAATNGYTPHLHISAREGQVDVASVLEAGA 519  
QY 518 VNVADLWKFTPLHEAAAKGYEICKILLQHGADPTKKNRDGNTPDLVKDGDIDQLDL 577  
Db 520 AHSLATKKGFTPLHVAAGYSLDVAKLLLORRAADSAGKNGLTPLHVAHYDN----- 573  
QY 578 RGDAAALDAAKGCCLARVKKLSSPDNVNCRDGTGRHSTPLHLAGYNNLEVAEYLLQHGA 637  
Db 574 -----QKVALLLLEKASP-----HATAKNGYTPHLHIAAKKNQMOIASTLLNYGA 618  
QY 638 DVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTOCAL 697  
Db 619 ETNTVTKQGYTPHLHSAOEGHTDMVTLLVLEKGANIHMTSGSLTSLHAAEDKVNADI 678  
QY 698 LLAHGADPTLKNQOGTDLIV-----SADDVSALLTAAMPSPA-LPSCYKP----- 743  
Db 679 LTKHGADODATKLGTYPLIVACHYGNVKNVFLKQGANVNAKTKNGYTPHLHQAQOQH 738  
QY 744 -OVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASLEK 802  
Db 739 THIIINVLLQHGAKPNATTA--NGNTALAIARLGYI--SVVDTLKVVVTEVTTTTTITE 794  
QY 803 KEVPGVDSITOFVRLGLEHLMDFERE---QITLDVLVEMGHKELKEIG 850  
Db 795 KHLNAPETMT-----VLVDSEEGDDTMTGGEYLRPEDLKELG 836

## RESULT 7

US-09-031-485-33  
; Sequence 33, Application US/09031485  
; Patent No. 5824306  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Liang  
; APPLICANT: Blehm, E. Scot  
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCE ADDRESSES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; ADDRESS: Heskia Corporation  
; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: WordPerfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/031.485  
; FILING DATE:

```

Db      646  TPLHAAQEGHREMAALLIENGAKVGQAQNRGLTPMHLCAQEDRVSAEELVKENAIDP 705
QY      735  SALP-----SCKPQVLNVR---SPGATADALSGSPSPSLSAASSLDLNGSFSE 784
Db      706  KTKAGYTPLHVACHFGQT-NMVRFLIEHGARVSIVTRASYTPLHQAAQQG-----754
QY      785  LSSVSSSGTGGASSLEKKKPGVDFSTQ-----FVRNLG----LEHLMDIFEREQIT-L 835
Db      755  HNSVRYLLEHGAS-----PNVHTSTGTPTLSIAERLGYVSVVEALKTITETTIVTIT 807
QY      836  DVLVMGHK-----ELKEIGINAYGH--RHKLIKGV 864
Db      808  TIVTEERYKQPONPEAMNETMFSDSEGEDNQITANAHADFSFSLTKGL 857

RESULT 8
US -08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US -08-847-429A-33

Query Match 14.2%; Score 821; DB 2; Length 1745;
Best Local similarity 29.1%; Pred. No. 5.6e-65;
Matches 276; Conservative 125; Mismatches 321; Indels 228; Gaps

QY      3  GRKDVVEYLIONGASVQARDGGLPLHNACSFGEAEVVVNLRLRHGADPNARDNWNVTPL 62
Db      48  GNLDVRVLELRSGTDINTNAGLNALHLSKEGHEVYVRELLRKADVDAAATKRGNTAL 107
QY      63  HEAAIKGKIDVCIVLIHOHAETPIRNTDGRALDLADP-----SAKAVLTGE 109
Db      108  HIASLAGQLIVTVLVENGANVVSOLNGFTPLYMAAENHESVRYLLAHANAQALUSTE 167
QY      110  YKDELLESARGSNEEKKMMALITPLNVNCHASDGKRKSTP-LHLAGNRYNRKIVQLLQ-- 166
Db      168  DGFPTPLAVALOQCHDRVAVLLE-----NDRGKVRPLALHIAAKKDDTKAATLILONE 221

```

[illegible]

## RESULTS

RESULTS 3  
US-09-065-474-33

US-09-003-474-33  
: Sequence 33 Application IIS/09065474

: Sequence 33, Appl. No. 6063599

Patent No. 6063599  
GENERAL INFORMATION:

GENERAL INFORMATION:

APPLICANT: Tang, Liang  
APPLICANT: F. Scott

APPLICANT: BIEHM, E. SCOT

; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
; TITLE OF INVENTION: PROTEINS NUCLEIC ACID MOLECULES AND

**TITLE OF INVENTION:** PROTEINS, NUCLEIC ACID MOLECULES, AND  
USES THEREOF

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

; NUMBER OF SEQUENCES: 1  
CORRELATION COEFFICIENT: .9876

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carol Talkin

; ADDRESSEE: Heska Corporation

STREET: 1825 Sharp

; CITY: Fort Collins

; STATE: Color

; COUNTRY: USA

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-065-474-33

```

Query Match	14.28	Score 821	DB 3	Length 1745
-------------	-------	-----------	------	-------------

Query Match

BEST LOCAL SIMILARITY	29.1%	F1ED: NO: 3.00	33%
Matches	276:	Conservative	125:
Mismatches	321:	Indels	228:
Gaps	28:	Gaps	28:

Qy	3	GRKDVEYLLONGASVQARDGGLTPLHNACSFGEAEVYVNLRLRGADPNARDNWNTPPL	62
Db	48	GNLDRLVLELRSGTDINTCNANGNLNALHVASKEGHEHVYRELLKKRDAVDAATRKGN TAL	107
Qy	63	HEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADP-----SAKAVLTGE	109
Db	108	HIASLAGOELIVTVLVEANGANNVQSLNGFTPLYMAAQENHESVVYRLLAHNAQALSTE	167
Qy	110	YKDELLESARSNEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYRVKVIQVLLQO--	166
Db	168	DGFTPLAVALQOQGHDRVAVILLE-----NDTRGVRLPALHIAAKKDDTKAATLLQNE	221
Qy	167	HGADVHAKDKGLVPLHNACSYGHEVTELLVKKHGCYVAMDLWFOFTPLHEAASKNRVEV	226
Db	222	HNSDV--TSKSGFTPLHIAAHYGENVAQVLLLEKGANVYQARHNISPLHVATKWGRTNM	279
Qy	227	CSLLSYGADPTLINCNKSAIDLPTPQLKERLAYEFKGHSLQQAAREADVTRIKKHL	286
Db	280	VSLLLAHG---VIDCTRROLL-----	298
Qy	287	LEMVNFKIPQTHETALHCAASAPYPRKQICELLLRKGANINEKTEFUTPLHVASEKAH	346
Db	299	-----TPLHCASRSGH---DQVVDLLEKGAIPISAKTKNGLAPLHMAAQ---	339
Qy	347	NDVVEVVVKHEAKYNALDNLGOSTLHRAAYCGHLQTCRLLLSYCGDPNTISLQGTFTALQ	406
Db	340	--VDDVTVDY-----LTPLHVAAHCGHVRAKLLLDNRADPNARALNGFTPLHI	386
Qy	407	G-----NENVQOLLQEGIS-----LGNSEADROLLEAAKAGDVETVKK	444
Db	387	ACKNRIKIVELLIKYHAAIEATETESGLSPLHVAAFMGAINIVYLLQOGANADVATV--	444
Qy	445	LCTQSVNCRDIEGRQSTPLHFAAGYNRVSVYLLQHGADVHAKDKGGLVPLVHNACSYG	504
Db	445	-----RGETPLHAAARANQTDIVRVLPNGAQVDAARELOQTPLHIASRLG	490
Qy	505	HYEAEVLVKHGVVNVADLWKFTPLHAAAKKYETCKLLLOHQGADPTKKNRDGNVTPLD	564
Db	491	NTDIVILLQANASPNAASTRDLYTPLHIAAKEGEEVAAITLMDHGDTKTLTKKGGFTPLH	550
Qy	565	LVKDGDTDIQLDKCDRAALLDAAKKGCLARVKKL--SSPDNVNCRDTCGRHS-TPLHLA	620
Db	551	L-----AAKYGNLPVAKSLLERGT-----VDIEGKNQVTFPLHVA	585
Qy	621	AGYNNLEVAEYLLQHGADVNADCKGLTPLHNAASYGHDVDAALLIKYKNCVNAATDKAF	680

```

Db 586 AHYNNDKVALLENGASAAHAAKNGYTPHLIAAKNQMDIASTLLHYKANANAESKAGF 645
QY 681 TPLHEAAQKGRTOICALLLAHAGADPTLKNOEGOTPLDLVSADDSALL-----TAAMPP 734
Db 646 TPLHAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSAEELVKENAAIDP 705
QY 735 SALP-----SCYKPOVLNGVR---SPGATADALSSGSPSSLSAASSLDNLGSGFSE 784
Db 706 KTKAGYTPLVACHFGQI-NMVRFLIEHGARGVSVITRASYPHLHQAQOG-----754
QY 785 LSSVSSSGTEGASLEKKEVPGVDFSITQ-----FVRNLG-----LEHLMDFEREQIT-L 835
Db 755 HNSVRYLLEHGAS-----PNVHTSGOTPLSLAEELGVSVVEALKTITETTVITET 807
QY 836 DVLVEMGHK-----ELKEIGINAYGH--RHKLKGV 864
Db 808 TTVTEERYKQPONPEAMNETMFSDEDEGDNQITANAHADFSESLTKGL 857

RESULT 10
US-09-557-034-33
; Sequence 33, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-557-034-33

```

```

Query Match 14.2%; Score 821; DB 4; Length 1745;
Best Local Similarity 29.1%; Pred. No. 5,6e-65;
Matches 276; Conservative 125; Mismatches 321; Indels 228; Gaps 28;
QY 3 GRKDVVEYLQNGASVQARDGGLTPLHNACSFGEAEVVVNLRLRGADPNARDNNNYTPL 62

```

```

Db 48 GNLDRLVLELLRSRGTINTCNANGLNALHLASKEGHHVVRELLKRKADYDAATRKNATL 107
QY 63 HEAAIKGKIDVICVILQHGAEPTIRNTDGTALDLADP-----SAKAVLTGE 109
Db 108 HIASLAGQELIVTLVENGANVNSQNGFTPLYMAAQENHESVVRVYLLAHANANQALSTE 167
QY 110 YKKDELLSARSNGNEKEMWALLTPLNVNCHASDGKRKSTP-LHLAAGYNRVKIVQLLLQ-- 166
Db 168 DGFTPLAVALQOQHDRVAVLLE-----NDTRGVRLPALHIAAKKDDTKAATLLQNE 221
QY 167 HGADVHAKDKGDLVPLHNACSYGHVETELVKHGACVYVNDLQWOTPLHEAASKNRVVE 226
Db 222 HNSDV--TSKSGFTPLHAAHGNVENVAQLLEKGANVYQARHNISPLHVATKNGRTNM 279
QY 227 CSLLSYSGADPTLLNCHNKSAIDLAPTQLKERLAYEFKGHSLLOAAREADVTRIKKHL 286
Db 280 VSLLLAHGA--VIDCTRDL--298
QY 287 LEMVNFKHPQTHETALHCAAAAPYPRKQICELLRLKGANINEKTEFELPLHVAASEKAH 346
Db 299 -----TPLHCASRSGH---DQVVDLLEKGAIPISAKTKGLPLHMAA---339
QY 347 NDVVEVVVHKAQVNALDNLGQTSLHRAAYCGHLOTCRLLSYCGDPNIIISLOGETALQM 406
Db 340 --VDDVTVDY-----LTPLHVAACHGVHVKLLLDNRADPNARALNGFTPLHI 386
QY 407 G-----NENVOOLLQEGIS-----LQNSEADROLLPAAKAGDVETVKK 444
Db 387 ACKNRKIVELLKYLHAAIEATTESGLSPHVAAFMGAINIVIVYLQOGANADVTV--444
QY 445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVYLLQHGADVHAKDKGGLVPLHNACSYG 504
Db 445 -----RGETPLHAAARANOITDVRVLVRNGAQVDAARELOTPPLHIASRLG 490
QY 505 HYEVAELLVKHGAQVNVADLWKFPTLHEAAKGYEICKLLOLHOGADPTKKNRDGNTPLD 564
Db 491 NTDIVILLQANASPNAAATRDLYTPLHIAKEGQEEVAAIILMDHGTDKTLTKKGFTPLH 550
QY 565 LVKDGDTDIODLLRGDAALLDAKKGCLARVKL---SSPDVNCRTDQGRHS--TPLHLA 620
Db 551 L-----AAKYNLPAKSLLEGRTP-----VDIEGKNQVTPHLHVA 585
QY 621 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAAALLIKYCANVATDKWAF 680
Db 586 AHYNNDKVALLENGASAAHAAKNGYTPHLIAAKNQMDIASTLLHYKANANAESKAGF 645
QY 681 TPLHEAAQKGRTOICALLLAHAGADPTLKNOEGOTPLDLVSADDSALL-----TAAMPP 734
Db 646 TPLHAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSAEELVKENAAIDP 705
QY 735 SALP-----SCYKPOVLNGVR---SPGATADALSSGSPSSLSAASSLDNLGSGFSE 784
Db 706 KTKAGYTPLVACHFGQI-NMVRFLIEHGARGVSVITRASYPHLHQAQOG-----754
QY 785 LSSVSSSGTEGASLEKKEVPGVDFSITQ-----FVRNLG-----LEHLMDFEREQIT-L 835
Db 755 HNSVRYLLEHGAS-----PNVHTSGOTPLSLAEELGVSVVEALKTITETTVITET 807
QY 836 DVLVEMGHK-----ELKEIGINAYGH--RHKLKGV 864
Db 808 TTVTEERYKQPONPEAMNETMFSDEDEGDNQITANAHADFSESLTKGL 857

```

```

RESULT 11
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devatarajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identifi
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A

```

Wed Feb 12 16:08:51 2003

; CURRENT FILING DATE: 1998-05-21  
; EARLIER APPLICATION NUMBER: 60/047356  
; EARLIER FILING DATE: 1997-05-21  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1088  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-082-059-2

Query Match 10.9% Score 629.5; DB 4; Length 1088;  
Best Local Similarity 28.1%; Pred. No. 6.7e-48; Indels 181; Gaps 16;  
Matches 195; Conservative 85; Mismatches 234; Indels 181; Gaps 16;  
QY 164 LQHGADVHARDKGDVPLPHNACSYGHEYTELLVKGACVNMADLWQFTPLHEAASKNR 223  
DB 7 LLDKKNPNKALNGFTPLHACKNRIKVMELLKHGASIQAVTESGLTPHVAAPMGH 66  
QY 224 VEVCSLLSYGADPTLLCHNKSALIDAPTPQLKRLAYEFKHSILQAAREADVTRIKK 283  
DB 67 VNIVSQMLHGHGSPNTTVRG----- 87  
QY 284 HLSLEWVFKHPQHTALHCAAAAPYKPKKOICELLRLKGANINEKTEFLTPHVASE 343  
DB 88 -----ETALHHAARS---GQAEVYRLVQGAQVEAKAKQDQTPHISAR 129  
QY 344 KAHNDVVEVVKHEAKVNALNIGQTSUHRAYCGHLOTCLRLSYGCDPNIIISLOQFETA 403  
DB 130 LKADIVQQLLOQASPNAAATTSYTPPLHLSAREGHEDVAAFLLDHGASLSITTKKGF-- 187  
QY 404 LQMGNEVQQLLOQEGISLGNSEADRLLEAAKAGDVETVKLCVTQSVNCRDIEGRQSTP 463  
DB 188 -----TP 189  
QY 464 LHFAAGYNRVSVVEYLQHGADVHAKDKGGLVPLHNACSYGHEYAEILLVKGAVNVVAD 523  
DB 190 LHVAAYKYLEVANLLLOKSASPDRAKSGSLTPLVHAAHYDNQKVALLLLOQASPHAAA 249  
QY 524 LKWFPTLHEAAKGYEICKLLQHGADPTKKNRQNTPLDL-VKDGDTDQDILLRGDAA 582  
DB 250 KNGYTPHIAAKKNQMDIATTLLEYGADANAVTROGASVHLAOGFHVDMVSVLLGRNA 309  
QY 583 LLDAKKGCLARVKLLSPDNVNCRTQCRHSTPLHLAAGYNLEVAEYLLQHGADVNAQ 642  
DB 310 -----NVNLSNKG--LTPHLAAQEDRVNVAEVLNQGAVHDAQ 347  
QY 643 DKGGGLPLHNAASYGHVDVAAALLIKYNACVNATDKWAFPTLHEAAKQRTQCALLLAHG 702  
DB 348 TKMGYTPLVHVGCHYGNIKVNFLLQHSKVNKATKNGYTPHQAQQQGHITHIINVLLQNN 407  
QY 703 ADPTLKNQEGQTPDL-----VSADDVSALLT-AAMPSPALPSCKY---POVLNGV--R 750  
DB 408 ASPNELTVNGNTALGIARBLGYSVVDTLKIIVTEMTTTTTTEKHKNVPTETMEVLDM 467  
QY 751 SPQATADALSS-----GP-----SSPS-----SLSAASLDNLGSGFSLSVV 789  
DB 468 SDDEGEDAMTQDKYLGQDLKELGDDSLPAEGYMGFSLGARSASDR---SY-----TLN 520  
QY 790 SSSGTEGASLEKVEYGVDFSTQTVRNIGLEHL 824  
DB 521 RSSYARDSMMIELLVPSKEQHLT-FTREFDSDSL 554

## RESULT 12

US-09-065-474-139  
; Sequence 139, Application US/09065474  
; Patent No. 6063599  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Liang  
; APPLICANT: Blehm, E. Scot  
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 171  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; ADDRESSEE: Heskia Corporation  
; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: WordPerfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/065,474  
; FILING DATE: 24-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verser, Carol Talkington  
; REGISTRATION NUMBER: 37,459  
; REFERENCE/DOCKET NUMBER: HW-5-C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 970/493-7272  
; TELEFAX: 970/484-9505  
; INFORMATION FOR SEQ ID NO: 139:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 352 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
; US-09-065-474-139

Query Match 8.9% Score 513; DB 3; Length 352;

Best Local Similarity 30.2%; Pred. No. 3.7e-38; Indels 86; Gaps 5;  
Matches 127; Conservative 64; Mismatches 144; Indels 86; Gaps 5;  
QY 299 ETALHCAAAAPYKPKKOICELLRLKGANINEKTEFLTPHVASEKAHNDVVEVVKHEA 358  
DB 12 ETPLHLAARA---NQTDIVRVLVNAGVDAARLQTPHLHIAASRLGNTDIVILLQANA 68  
QY 359 KVNALDNLGQTSUHRAYCGHLOTCLRLSYGCDPNIIISLOQFETAQMGNEVQQLLOEG 418  
DB 69 SPNAATRDLYTPHLHIAAKEGQEEVAAILMDHGTDKTLTKKGFPLHL----- 116  
QY 419 ISLGNSEADRLLEAAKAGDVETVKLCVTQSVNCRDIEGR-QSTPLHFAAGYNRVSVVE 477  
DB 117 -----AAKYNLPVAKSL--LERGTPVDIEGKNQVTPLVHAAHYNNDKVAL 160  
QY 478 YLQHGADVHAKDKGGLVPLHNACSYGHEYAEILLVKGAVNVVADLWKFPTLHEAAAKG 537  
DB 161 LLENGASAAHAKNGYTPHLHIAAKKNQMDIATLLHYKANANAEKAGETPLHLAAOEG 220  
QY 538 KYEICKLLQHGADPTKKNRQNTPLDLVKGDDTDIODLLRGDAALLDAKKGCLARVK 597  
DB 221 HREMAALLIENGAKVGAQRNG----- 242  
QY 598 LSSPDNVNCRDTQGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLPLHNAASYG 657  
DB 243 -----LTPMHLCAQEDRVSAEELVKENAAIDPKTKAGYTPLVHACHFG 286  
QY 658 HVDVAALLIKYNACVNATDKWAFPTLHEAAKQRTQCALLLAHGADPTLKNQEGQTPDL 717  
DB 287 QINMYRFLIEHGAHSVITRASVYTPHQAQQQGHNSVVYRLLLEHGAASPNVHTSTGTG 346  
QY 718 L 718  
DB 347 I 347

## RESULT 13

US-09-557-034-139

1



Db 118 NOTDIVRVLRNGAQVDAARELOTPHLIASRLGNTDIVILLQANASPNATRDLYTPL 177  
QY 564 DL-VKDGDTDIDLLRG---DAALLDAAKKGLARVKLLSSPDNVNCRDQTGRHSTPLHL 619  
Db 178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG-----FTPLHL 212  
QY 620 AAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWA 679  
Db 213 AAKYGNLPVAKSLLEGTVPDIEGKNQVTPLVAAHYNNNDKVALLLLENGASAHAAKNG 272  
QY 680 FTPLHAAAKGRTOLCALLLAHAGDPTLKNQ 710  
Db 273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303

RESULT 15

US-08-847-429A-23  
; Sequence 23, Application US/08847429A  
; Patent No. 5827692  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Liang  
; APPLICANT: Blehm, E. Scot  
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
; TITLE OF INVENTION: USES THEREOF  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carol Talkington Verser, Ph.D.  
; ADDRESSEE: Heskia Corporation  
; STREET: 1825 Sharp Point Drive  
; CITY: Fort Collins  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80525  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: WordPerfect for Windows, Version 7.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/847,429A  
; FILING DATE: 24-APR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verser, Carol Talkington  
; REGISTRATION NUMBER: 37,459  
; REFERENCE/DOCKET NUMBER: HW-5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 970/493-7272  
; TELEFAX: 970/484-9505  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 303 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-847-429A-23

Query Match 7.3%; Score 421.5; DB 2; Length 303;  
Best Local Similarity 30.4%; Pred. No. 6.2e-30;  
Matches 119; Conservative 52; Mismatches 125; Indels 95; Gaps 8;  
QY 327 INEKTKEFLPLHVAEKAHNDVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLOTCRL 386  
Db 1 VDDVTVDYLTPLHV-----AAHCGHVRVAKLL 27  
QY 387 LSYGCDPNILSLOGFTALQMNENVOQLQEGISLGNSEADROLLEAKAGDVETVKLC 446  
Db 28 LDRNADPNARALNGFTPLHI-----ACKKNRIKIVLELL 61  
QY 447 TVQSVNCRDIEGROS---TPLHFAAGYNRVVVEYLLQHGADVHAKDKGLVPLHNACSY 503  
Db 62 KYHAA----IEATTESGLSLHVAAFMGAINIVYLLQOGANADVATVRGETPLHLAARA 117

QY 504 GHYEVAELLVKHCAVNVADLWKFTPLHGAAGKYEICKLLLOHGADPTKKNRDGNTP 563  
Db 118 NOTDIVRVLRNGAQVDAARELOTPHLIASRLGNTDIVILLQANASPNATRDLYTPL 177  
QY 564 DL-VKDGDTDIDLLRG---DAALLDAAKKGLARVKLLSSPDNVNCRDQTGRHSTPLHL 619  
Db 178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG-----FTPLHL 212  
QY 620 AAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWA 679  
Db 213 AAKYGNLPVAKSLLEGTVPDIEGKNQVTPLVAAHYNNNDKVALLLLENGASAHAAKNG 272  
QY 680 FTPLHAAAKGRTOLCALLLAHAGDPTLKNQ 710  
Db 273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303  
Search completed: February 12, 2003, 12:27:18  
Job time : 21.0427 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:33:09 ; Search time 22.5641 Seconds  
(without alignments)  
1245.508 Million cell updates/sec

Title: US-09-843-159b-3  
Perfect score: 5769  
Sequence: 1 GFRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPGMVDG 1100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5602	97.1	1074	10	US-09-509-196A-2
2	4889.5	84.8	1327	10	US-09-841-835-2
3	3596	62.3	802	9	US-09-964-899-41
4	3093	53.6	949	10	US-09-841-835-10
5	1918.5	33.3	673	10	US-09-841-835-8
6	798.5	13.8	1724	9	US-09-964-899-43
7	395	6.8	426	10	US-09-908-711-70
8	379.5	6.6	1719	9	US-10-012-896-378
9	379.5	6.6	1719	9	US-09-895-793-378
10	379.5	6.6	1719	9	US-09-895-814-378
11	379.5	6.6	1719	10	US-09-759-143-378
12	379.5	6.6	1719	10	US-09-780-669-378
13	379.5	6.6	1719	10	US-09-822-827-378
14	368.5	6.4	740	10	US-09-835-788A-12
15	332.5	5.8	835	10	US-09-947-199-2
16	324.5	5.6	835	10	US-09-947-199-8
17	312.5	5.4	599	10	US-09-735-368-2
18	297	5.1	551	10	US-09-835-788A-17
19	281.5	4.9	285	10	US-09-835-788A-18

20	268	4.6	306	10	US-09-835-788A-19	Sequence 19, Appl
21	267	4.6	1054	10	US-09-798-042-87	Sequence 87, Appl
22	258	4.5	251	10	US-09-835-788A-13	Sequence 13, Appl
23	247.5	4.3	328	10	US-09-758-593A-11	Sequence 11, Appl
24	247.5	4.3	328	10	US-09-758-593A-12	Sequence 12, Appl
25	247.5	4.3	329	10	US-09-880-192-62	Sequence 62, Appl
26	247.5	4.3	329	10	US-09-758-593A-1	Sequence 1, Appl
27	243.5	4.2	2444	10	US-09-944-849-2	Sequence 2, Appl
28	243	4.2	752	10	US-09-927-180-2	Sequence 2, Appl
29	240	4.2	671	9	US-09-924-400-306	Sequence 306, App
30	240	4.2	671	9	US-10-012-896-380	Sequence 380, App
31	240	4.2	671	9	US-09-895-793-380	Sequence 380, App
32	240	4.2	671	9	US-09-895-814-380	Sequence 380, App
33	240	4.2	671	10	US-09-825-301-10	Sequence 10, Appl
34	240	4.2	671	10	US-09-759-143-380	Sequence 380, App
35	240	4.2	671	10	US-09-780-669-380	Sequence 380, App
36	240	4.2	671	10	US-09-810-936-306	Sequence 306, App
37	240	4.2	671	10	US-09-822-827-380	Sequence 380, App
38	240	4.2	671	10	US-09-429-755-306	Sequence 306, App
39	237.5	4.1	656	9	US-09-924-400-305	Sequence 305, App
40	237.5	4.1	656	9	US-10-012-896-379	Sequence 379, App
41	237.5	4.1	656	9	US-09-895-793-379	Sequence 379, App
42	237.5	4.1	656	9	US-09-895-814-379	Sequence 379, App
43	237.5	4.1	656	10	US-09-825-301-9	Sequence 9, Appl
44	237.5	4.1	656	10	US-09-759-143-379	Sequence 379, App
45	237.5	4.1	656	10	US-09-780-669-379	Sequence 379, App

## ALIGNMENTS

## RESULT 1

US-09-509-196A-2

; Sequence 2, Application US/09509196A

; Patent No. US20020037582A1

; GENERAL INFORMATION:

; APPLICANT: DALY, Roger J.

; APPLICANT: SUTHERLAND, Robert L.

; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling

; TITLE OF INVENTION: Proteins

; FILE REFERENCE: 1871-129

; CURRENT APPLICATION NUMBER: US/09/509,196A

; CURRENT FILING DATE: 2000-03-23

; PRIOR APPLICATION NUMBER: P09388

; PRIOR FILING DATE: 1997-09-23

; PRIOR APPLICATION NUMBER: PCT AU98/00795

; PRIOR FILING DATE: 1998-09-23

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1074

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-509-196A-2

Query Match	97.1%	Score 5602;	DB 10;	Length 1074;
Best Local Similarity	99.4%	Pred. No. 0;		
Matches 1068;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0;
QY	27	IPLNACSFHAEVNVLLRLRGADPNARDNNWYTPLEHAAIKGKIDVCIVLLQHGAEPTI	86	
Db	1	IPLNACSFHAEVNVLLRLRGADPNARDNNWYTPLEHAAIKGKIDVCIVLLQHGAEPTI	60	
QY	87	RNTDRTALDIADPSAKAVLTGEYKKDELLESARGSEKEMWALLTPLNVNCHASDGRKS	146	
Db	61	RNTDRTALDIADPSAKAVLTGEYKKDELLESARGSEKEMWALLTPLNVNCHASDGRKS	120	
QY	147	TPHLAAGYNRVKIVQLLLQHGADVHAKDGLVPLHNACSYGVHYETELIVKKGACVNA	206	
Db	121	TPHLAAGYNRVKIVQLLLQHGADVHAKDGLVPLHNACSYGVHYETELIVKKGACVNA	180	
QY	207	MDLWQFTPLHEAASKNRVEVCSLLSYGADPTLLNCHNKSATDIAPTPOLKERLAYEFKG	266	

Db 181 MDLWQFTPLHEAASKNRVEVCSLLSYGADPTLLNCKNSAIDLAPTQPKERLAYEFKG 240  
QY HSLQAAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAASPYPKRKQICELLRRKAN 326  
Db 241 HSLQAAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAASPYPKRKQICELLRRKAN 300  
QY INEKTKEFTPLHVASEKAHNDVVEVVKHAEKVNALDNLGQTSLHRAAYCGHLQTCRLL 386  
Db 301 INEKTKEFTPLHVASEKAHNDVVEVVKHAEKVNALDNLGQTSLHRAAYCGHLQTCRLL 360  
QY LSYGCDPNIISLOGFTALQMGNEVQVLLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 446  
Db 361 LSYGCDPNIISLOGFTALQMGNEVQVLLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 420  
QY TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 506  
Db 421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480  
QY EVAELLVKGVAVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDCNTPLDLV 566  
Db 481 EVAELLVKGVAVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDCNTPLDLV 540  
QY KDGDTDQDLRLGRDAALLDAAGKGLARVKKLSPPDNVNCRDQTGRHSTPLHLAAGYNL 626  
Db 541 KDGDTDQDLRLGRDAALLDAAGKGLARVKKLSPPDNVNCRDQTGRHSTPLHLAAGYNL 600  
QY EVAEYLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEA 686  
Db 601 EVAEYLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNASLNTATDKWAFTPLHEA 660  
QY AQKGRQTCALLAHAGADPTLKNQEGOTPLDLSADDDVSALLTAAMPSPALPSCYKQV 746  
Db 661 AQKGRQTCALLAHAGADPTLKNQEGOTPLDLSADDDVSALLTAAMPSPALPSCYKQV 720  
QY NGVRSFGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASSLEKKEVP 806  
Db 721 NGVRSFGATADALSSGSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASSLEKKEVP 780  
QY GVDFTSQFVRLNGLBHLMDIFEREQITLDLVEMGHKELKEIGINAYGHRHKLKIGVER 866  
Db 781 GVDFTSQFVRLNGLBHLMDIFEREQITLDLVEMGHKELKEIGINAYGHRHKLKIGVER 840  
QY LISGOQGLNPYLT/LNTSGSGTILIDLSPDDKEFQSVSEEMQSTVRHRRDGGHAGGIFNRY 926  
Db 841 LISGOQGLNPYLT/LNTSGSGTILIDLSPDDKEFQSVSEEMQSTVRHRRDGGHAGGIFNRY 900  
QY NILKIQKVCNKKLWERYTHRRKEVSEENHNANERMLFHGSPFVNALIHKGFDERHAYIG 986  
Db 901 NILKIQKVCNKKLWERYTHRRKEVSEENHNANERMLFHGSPFVNALIHKGFDERHAYIG 960  
QY GMFGAGIYFAENSSKSNQVYIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSA 1046  
Db 961 GMFGAGIYFAENSSKSNQVYIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSA 1020  
QY MKMAHSPGCHSHVTRGSPVNGLALAEVLYVRGEQAYPEYLYIYQIMRPGMVDG 1100  
Db 1021 MKMAHSPGCHSHVTRGSPVNGLALAEVLYVRGEQAYPEYLYIYQIMRPGMVDG 1074

## RESULT 2

US-09-841-835-2

; Sequence 2, Application US/09841835

; Patent No. US20020076795A1

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber &amp; Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCET NUMBER: 600-1-230 CIPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1327 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-09-841-835-2

## Query Match

84.8%; Score 4889.5; DB 10; Length 1327;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

QY 1 GFGKDKVVEYLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNY 60  
Db 225 GFGKDKVVEYLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNY 284  
QY 61 PLHEAAIKGKIDVCIVLLOHGAEPTRNTDGRDALDAPSAKAVLTGEYKKDELLESA 120  
Db 285 PLHEAAIKGKIDVCIVLLOHGAEPTRNTDGRDALDAPSAKAVLTGEYKKDELLESA 344  
QY 121 SGNEEKMAALLPLNVNCHASDGRKSTPLHLAAGYNRVIVLQHGADVHAKDKGLV 180  
Db 345 SGNEEKMAALLPLNVNCHASDGRKSTPLHLAAGYNRVIVLQHGADVHAKDKGLV 404  
QY 181 PLHNACSYGHEVTELLVKGACVNAAMLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 240  
Db 405 PLHNACSYGHEVTELLVKGACVNAAMLWQFTPLHEAASKNRVEVCSLLSYGADPTLL 464  
QY 241 NCHNSAIDLAPTQPKERLAYEFKGHSLQAAAREADVTRIKKHLSLEMVNFKHPQTH 300  
Db 465 NCHNSAIDLAPTQPKERLAYEFKGHSLQAAAREADVTRIKKHLSLEMVNFKHPQTH 524  
QY 301 ALHCAASPYPKRKQICELLRRKANINEKTEFTPLHVASEKAHNDVVEVVKHAEK 360  
Db 525 ALHCAASPYPKRKQICELLRRKANINEKTEFTPLHVASEKAHNDVVEVVKHAEK 584  
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLOGFTALQMGNEVQVLLQEG 420  
Db 585 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLOGFTALQMGNEVQVLLQEG 644  
QY 421 LGNSEADROLLLEAAKAGDVETVKKLTCTVSVCNCRDIEGRQSTPLHFAAGYNRVSV 480  
Db 645 IRTSDVDYRLLLEAAKAGDVETVKKLTCTVSVCNCRDIEGRQSTPLHFAAGYNRVSV 704  
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVTELLVKGACVNAAMLWQFTPLHEAASKNR 540  
Db 705 QHGADVHAKDKGGLVPLHNACSYGHEVTELLVKGACVNAAMLWQFTPLHEAASKNR 764

```
QY 541 ICKLLQHGADPTKKNRDNTPDLVDKGDPTDIDLLRGDAALLDAAKKGCGLARVKKLS 600
Db 765 ICKLLKKGADPTKKNRDNTPDLVDKGDPTDIDLLRGDAALLDAAKKGCGLARVKKLCT 824
QY 601 PDNVCNCRDTQGRHSTPLHLAAGYNVAEYLLQHGADVNADQKGGTLPPLHNAASYGHVD 660
Db 825 PENINCRTQGRNSTPLHLAAGYNVAEYLLQHGADVNADQKGGTLPPLHNAASYGHVD 884
QY 661 VAAALLIKYNACVNATDKWAFPLHEAAQKGRTOICALLAHGADPTLKNQOGPTDLVS 720
Db 885 IAALLIKYNTCVNATDKWAFPLHEAAQKGRTOICALLAHGADPTMKNQOGPTDLAT 944
QY 721 ADVSALLTAAMPSPALSCYKPVNLGVSPGATADALSSGSPSSLSAASIDNLG 780
Db 945 ADDIRALLIDAMPPEALPTCKPQAT-----VVSASLISPASTPSCLSAASIDNLG 997
QY 781 SPSESSSVVSSGTEGASLSKK--EVPGVDFSTQFVNRNLGLEHLMDFIFEREQITLDVL 838
Db 998 PLAELAVGASNAGDGAAGTERKEGEVAGLDWNIISOFLKSLGLEHLRDFIFETEITLDVL 1057
QY 839 VEMGHKELKEIGINAYGHRHKLIGVERLISGOOGLNPLYLTNTSGSGTILIDLSPODK 898
Db 1058 ADMGHEULKEIGINAYGHRHKLIGVERLISGOOGLNPLYLTNTSGSGTILIDLSPEDK 1117
QY 899 FOSVSEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEENHHA 958
Db 1118 YQSVEEMQSTVREHRDGGNAGGIFNRYNIRIQVNVNKKLRERFCHQKEVSEENHNNH 1177
QY 959 NERMLFPHGSPFNVAIIHGKGFDRHAYIGMGFAGITYFAENSCKSNQYVYGGTGCVPV 1018
Db 1178 NERMLFPHGSPFNVAIIHGKGFDRHAYIGMGFAGITYFAENSCKSNQYVYGGTGCPTH 1237
QY 1019 KDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEVYVRG 1078
Db 1238 KDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEVYVRG 1297
QY 1079 EQAYPEYLITYQIMRPE 1095
Db 1298 EQAYPEYLITYQIMKPE 1314

RESULT 3
US-09-964-899-41
; Sequence 41, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-41

Query Match 62.3%; Score 3596; DB 9; Length 802;
Best Local Similarity 65.1%; Pred. No. 2.3e-253;
Matches 720; Conservative 5; Mismatches 5; Indels 376; Gaps 5;

QY 1 GFGRKDVVEYLLQNGASVQARDGGIPLHNAACFSGHAEVNVNLLRHGADPNARDNWNVT 60
Db 67 GFGRKDVVEYLLQNGANVQARDGGIPLHNAACFSGHAEVNVNLLRHGADPNARDNWNVT 126
QY 61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGR TALDLADPSAKAVLTGE-----YKKDEL 115
```

```
Db 127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGR TALDLADPSAKAVLTGRKSVSGYSRKPV 186
QY 116 LES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 174
Db 187 KNNLARGSGNEEKMMALLTPLNVNCHASDGR----- 216
QY 175 DKGDVPLHNACSYGHEVTELLVKGACVYNAMDLQWFTPLHEAASKNRVEVCSLLLSYG 234
Db 217 -----KHGACVYNAMDLQWFTPLHEAASKNRVEVCSLLLSYG 252
QY 235 ADPTLLMCHNKSALDAPTQPKERLAYEPKGHSLLOAAREADVTRIKKHLSELMVNFKH 294
Db 253 ADPTLLMCHNKSALDAPTQPKERLAYEPKGHSLLOAAREADVTRIKKHLSELMVNFKH 312
QY 295 POTHETALHCAASAPYKPKQICELLLRKCANINEKTEFLTPLHVASEKAHDVVVEVV 354
Db 313 POTHETAL----- 320
QY 355 KHEAKVNALDNLGOTSLHRAAYCGHLQTCRLLLSYCGDPNIIISLOGFTALQMGNEVQOL 414
Db 321 ----- 320
QY 415 LOEGISLGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRV 474
Db 321 -----KLCTVQSVNCRDIEGRQSTPLHFAAGYNRV 351
QY 475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGAVNVNADLWKTPTLHEAA 534
Db 352 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKHGAVNVNADLWKTPTLHEAA 411
QY 535 AKGYEICKLLOHGADPTKKNRDNTPDLVDKGDPTDIDLLRGDAALLDAAKKGCGLAR 594
Db 412 AKGYEICKLLOHGADPTKKNRDNTPDLVDKGDPTDIDLLRGDAALLDAAKKGCGLAR 471
QY 595 VKLSSPDNVCNCRDTQGRHSTPLHLAAGYNVAEYLLQHGADVNADQKGGTLPPLHNA 654
Db 472 VKLSSPDNVCNCRDTQGRHSTPLHLAAGYNVAEYLLQHGADVNADQKGGTLPPLHNA 531
QY 655 SYGHVDVAAALLIKYNACVNATDKWAFPLHEAAOKGRTOICALLAHGADPTLKNQOGPT 714
Db 532 SYG----- 534
QY 715 PLDLVSADVSALLTAAMPSPALSCYKPVNLGVSPGATADALSSGSPSSLSAAS 774
Db 535 ----- 534
QY 775 LDNLGSESELSSVVSSSGTEGASLSKKYVPYDFSTQFVNRNLGLEHLMDFIFEREQIT 834
Db 535 -----IT 536
QY 835 LDVLVEMGHKELKEIGINAYGHRHKLIGVERLISGOOGLNPLYLTNTSGSGTILIDLS 894
Db 537 LDVLVEMGHKELKEIGINAYGHRHKLIGVERLISGOOGLNPLYLTNTSGSGTILIDLS 596
QY 895 DDKEFQSVSEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEEN 954
Db 597 DDKEFQSVSEEMQSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERYTHRRKEVSEEN 656
QY 955 HNANERMLFPHGSPFNVAIIHGKGFDRHAYIGMGFAGITYFAENSCKSNQYVYGGTGC 1014
Db 657 HNANERMLFPHGSPFNVAIIHGKGFDRHAYIGMGFAGITYFAENSCKSNQYVYGGTGC 716
QY 1015 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEV 1074
Db 717 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEV 776
QY 1075 IYRGSQAYPEYLITYQIMRPEGMVDG 1100
Db 777 IYRGSQAYPEYLITYQIMRPEGMVDG 802

RESULT 4
```

US-09-841-835-10  
; Sequence 10, Application US/09841835  
; Patent No. US20020076795A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 949 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-841-835-10

Query Match 53.6%; Score 3093; DB 10; Length 949;  
Best Local Similarity 81.5%; Pred. No. 1.1e-216;  
Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;

QY 1 GFGKDVVEYLLQNGASVOARDGGGLPLHNACSFGEAEVWVLLLRHGDADPNARDNNYT 60  
DB 225 GFGKDVVEYLLQNGANVANRDGGGLPLHNACSFGEAEVWVLLLRHGDADPNARDNNYT 284  
QY 61 PLHEAAIKGIDVICVLLQHGAEPTIRNTDGRALDADPSAKAVLTGEYKDELLESAR 120  
DB 285 PLHEAAIKGIDVICVLLQHGADPNRNTDGSALDADPSAKAVLTGEYKDELLEAAR 344  
QY 121 SGNEEKWALLPLNPNCHASGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 180  
DB 345 SGNEEKWALLPLNPNCHASGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 404  
QY 181 PLHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAAKRNVEVCSLLLSYGADPTLL 240  
DB 405 PLHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAAKRNVEVCSLLLSYGADPTLV 464  
QY 241 NCHNSAIDLAPTPOLKERLAYEFKSHLSLQAAAREADVTRIKKHLSEMVNFKHPOTHE 300  
DB 465 NCHGSAVDMAPTPELRRLTYEFKSHLSLQAAAREADLAKVKTALALEINFKQFSHET 524  
QY 301 ALHCAASAPYPRKQICELLRLKGANINKEFTLTPHVASEKAHNDVVEVVKHEAV 360  
DB 525 ALHCAVASLHPKQVTELLLRKANVNEKNKDFMTPLHVAAERAHNDVVEVLRKHGAKM 584  
QY 361 NALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISLQGFTALQMGNEVQQLQEGIS 420

DB 585 NALDTLQOTALHRAALAGHLQTCRLLSYGSDPSTIISLQGTAAQMGNEAVQQLSESTP 644  
QY 421 LGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480  
DB 645 IRTSDVDYRLLEASKAGDLETVKOLCSSQNVNCRDIEGRHSTPLHFAAGYNRVSVVEYLL 704  
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVVAELLYKHGAVVNVADLWKFTPLHEAAKAGYE 540  
DB 705 HHGADVHAKDKGGLVPLHNACSYGHEVVAELLYKHGASVNVADLWKFTPLHEAAKAGYE 764  
QY 541 ICKLLQHGADPTKKNRDGNTPLDLVRDGDPTDIDLLRGGDAALLDAAKKGCLARVKLLSS 600  
DB 765 ICKLLKHGADPTKKNRDGNTPLDLVRDGDPTDIDLLRGGDAALLDAAKKGCLARVKLCT 824  
QY 601 PDVNCRTQGRHSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH- 658  
DB 825 PENINCRDTQGRNSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGCL 884  
QY 659 VDVAALLIKYNACVNTDKWAFPLHEAAOKGRQTOLCALLLHAGADPTLNQEGQTPL 716  
DB 885 ARVOKLCTPENINCRTQGRNSTPLHLAAGYNLNLEVAEYLLQHGADVNAQDKGGLIPL 942

## RESULT 5

US-09-841-835-8  
; Sequence 8, Application US/09841835  
; Patent No. US20020076795A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 673 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-841-835-8

Query Match 33.3%; Score 1918.5; DB 10; Length 673;  
Best Local Similarity 80.9%; Pred. No. 1.7e-131;  
Matches 364; Conservative 42; Mismatches 43; Indels 1; Gaps 1;

QY 1 GFGKDVVEYLLONGASVOARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 60  
D5 225 GFGKDVVEYLLONGASVHARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYT 284  
QY 61 PLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLEAAR 120  
D5 285 PLHEAAIKGIDVCIVLLQHGADPNIRNTDGRSALDADPSAKAVLTGEYKKDELLEAAR 344  
QY 121 SONEKMMALLTPLNVNCHASGRKSTPLHLAAGYNNRVKIVOLLQHGADVHAKDGLV 180  
D5 345 SONEKMMALLTPLNVNCHASGRKSTPLHLAAGYNNRVKIVOLLQHGADVHAKDGLV 404  
QY 181 PLHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAASKNRVEVCSLLSYGADPTLL 240  
D5 405 PLHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAASKNRVEVCSLLSYGADPTLL 464  
QY 241 NCHNKSALDAPTQPKERLAYEFGHSLLOAAREADYTRIKKHLSEMVNFKHPQTHET 300  
D5 465 NCHGKSAMDAPTPELRRLTYEFGHSLLOAAREADYTRIKKHLSEMVNFKHPQTHET 524  
QY 301 ALHCAAASPYPRKQICELLRKGANINEKTEFLTPHVAASEKAHNDVVEVYVYKHEAKV 360  
D5 525 ALHCAVASLHPRKQVETELVKGACVAMDLWOFTPLHEAASKNRVEVCSLLSYGADPTLL 584  
QY 361 NALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDPNIISLOGFTALQWGNVQQLQEGIS 420  
D5 585 NALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDPNIISLOGFTALQWGNVQQLQEGIS 644  
QY 421 LGNSEADROLLEAAKAGDVETVKKLTCTVOS 450  
D5 645 SDPSIISLOGFTAAQMGN-EAVQQLSGHS 673

RESULT 6

US-09-964-899-43  
; Sequence 43, Application US/09964899  
; Patent No. US2002017446A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Dalia et al.  
; TITLE OF INVENTION: Identification of Genes Involved in  
; FILE REFERENCE: 4-31612 A  
; CURRENT APPLICATION NUMBER: US/09/964,899  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/236,893  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/298,309  
; PRIOR FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 1724  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-964-899-43

Query Match 13.8%; Score 798.5; DB 9; Length 1724;  
Best Local Similarity 25.2%; Pred. No. 1.9e-49;  
Matches 279; Conservative 167; Mismatches 380; Indels 279; Gaps 28;  
QY 3 GRKDVVEYLLONGASVOARDGGLPLHNACSFHAEVNVNLLRHGADPNARDNNYTPL 62  
D5 47 GHVEVSELLQREANVDAATKKGNTALHSLAGAEVVKVLTNGANVNAQSQNGFTPL 106  
QY 63 HEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAVLTGEYKKDELLEAAR 122  
D5 107 YMAAQENHLEVYKFLDNGASQSLSATGEGTPLAVA-----LQQG 146  
QY 123 NEEKMMALLTPLNVNCHASGRKSTPLHLAAGYNNRVKIVOLLQHGADVHAKDGLV 181  
D5 147 HDQVVSLLLE-----NDTKGVKPLPALHTAARKDDYTKAALLQNDNNADVESKSGFT 200  
QY 182 LHNACSYGHEVTELLVKGACVAMDLWOFTPLHEAASKNRVEVCSLLSYGADPTLLN 241

D5 201 LHIAAHYGNINVAATLLNRAAAVDFTARNITPLHVASKRGANMYKLLLDLDRGA----- 254  
QY 242 CHNKSALDAPTQPKERLAYEFGHSLLOAAREADYTRIKKHLSEMVNFKHPQTHETA 301  
D5 255 -----KIDATROGL-----TP 266  
QY 302 LHCAAASPYPRKQICELLRKGANINEKTEFLTPHVAASEKAHNDVVEVYVYKHEAKV 361  
D5 267 LHCGARSCH---EQVEMLLDRAAPILSKTKNGLSPLHMAATQGDHLCNVQQLLQHNVPVD 323  
QY 362 ALDNLGOTSLHRAAYCGHLOTCTRLLSYGCDPNIISLOGFTALQWGNVQQLQEGIS 421  
D5 324 DVTNDYLTALHVAACHGHYKAVKVLDDKANKANAKALNGFTPLHI----- 368  
QY 422 GNSEADROLLEAAKAGDVETVKKLTCTVOSNCRDIEGRQSTPLHFAAGYNNRVSVVEYLQ 481  
D5 369 -----ACKKNRIK-VMEALLKHGASIQAVTERGETALHMAARSQAQVVRVLYQ 416  
QY 482 HGADVHAADKGLVPLHNACSYGHEVYAEVALLVKGACVAMDLWOFTPLHVAASEKAHNDVVEVYVYKHEAKV 541  
D5 417 DGAQVEAKAKDDQTPHLHSAIRLKGADIVQQLQOQASPNAAATTSYGTPLHLSAREGHEDV 476  
QY 542 CKLLLOHGADPTKKNRDNTPDLVKGDDTDIOQLLRGDAALLD-----AAKKGCLAR 594  
D5 477 AAFLLDHGASLSITTKSGLTPLHVAHYDNOKVALL-----LDDQASPHAAKNG----- 527  
QY 595 VKLSSPDNVNCRDTQGRHSTPLHAAAGYNNVAEYLLQHGADVNAQDQGLPLHNA 654  
D5 528 -----YTPHLIAAKKNQMDIATLLEYGADANAVTROGIAVHLAA 568  
QY 655 SYGHVDVAALLIKYNACVNAATDKWAFTPLEAQAOKGRTQLCALLAHGADPTLNQ----- 710  
D5 569 QEGHVDVMSLLGRNANVNLNKSGLTPLHAAQEDRVNVAEVLVNOGAHVDAQTQKVGYP 628  
QY 711 ---EGOTPLDLYSADDVSALLTAAMPSPALPCYKQVNLGVNVRSPGATADALSSGSPSS 767  
D5 629 PLPHGKECVHLVTANGYTPHQAQQOQHT-----HIINVLQNNASPNELTVTVTEKH 681  
QY 768 SLASAASS-----LDNLSSGFSSELSSVSSGTEGASSLEK-----KEVPGV-DFSITQFV 816  
D5 682 KMNVEPETHNEVLNDSDDVEVRKANAPMELSDGEYISDVEEGNRCWTYKIPKVOEFTVK--- 738  
QY 817 RNLGLEHMLDIPEREQITVDLIVE-----MGHKLKEIGINAY----- 854  
D5 739 -----TDTFKREAFDVGLISTSAGEDAMTGDYKLGPODLKELGDDSLPABGYMGF 790  
QY 855 -----GHRHKLKQVE-----RLISQOQGLNPLY 878  
D5 791 SLGARSARFLVSFVMDARGGSMRGRSHHGMRIIPRCKTAPTRITCRVLRKHLANPPP 850  
QY 879 TLNTSGSGTILIDLSPPDKEFOS-VEEEMQSTVREHROGGHAGIFNRYNLIKQKVCNK 937  
D5 851 MVEGELASRLVEMGMPAGAPLGPVIVEIP-----HFGSMRGRERELIVLRSNG 900  
QY 938 KLWE--RYTHRRKEVSE-----ENHNHANERMLFHGSPFVNALHKGDFERHAYIGGFG 990  
D5 901 ETWKEHQFDSKNDUTELLNGDELDSPLEL---GKKRICRITITKDPQ----- 947  
QY 991 AGTYFAENS---SKSNQVYVIGGG 1012  
D5 948 ---YFAVWSRIKQESNQ---IGPEGG 967

RESULT 7

US-09-908-711-70  
; Sequence 70, Application US/09908711  
; Patent No. US20020045230A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA128  
; CURRENT APPLICATION NUMBER: US/09/908,711

```

; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE.:
; NAME/KEY: SITE
; LOCATION: (148)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (258)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (396)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (413)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (417)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (418)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (421)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; -S-09-908-711-70

```

Best Local Similarity 27.9%; Pred. No. 5e-21;

Matches 123; Conservative 52; Mismatches 1

[illegible]



Wed Feb 12 16:08:51 2003

Db 367 YLTALHVAACHGHHYKAKVLL 387

RESULT 8

US-10-012-896-378

; Sequence 378, Application US/10012896

; Publication No. US20020183251A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedrick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Meagher, Madeleine Joy

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C27

; CURRENT APPLICATION NUMBER: US/10/012,896

; CURRENT FILING DATE: 2001-12-10

; NUMBER OF SEQ ID NOS: 1011

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 378

; LENGTH: 1719

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-012-896-378

Query Match 6.6%; Score 379.5; DB 9; Length 1719;

Best Local Similarity 20.9%; Pred. No. 5.6e-19;

Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

QY 143 GRKSTPLHAAAGYNNRKYIYVOLLQ-HGADVHAKDKGDLVPLHNAACSYGHYVTELLVKG 201

Db 137 GEDLDKLHRAAMWGKVPKDLIVMLRDTDVNKKDKOKRTALHLASANGSEVVKLLDDR 196

QY 202 ACVNAMDLWQFTPLHAAASNNRVEVCSLLSYGADPTLLNCHNKSALDIAPTPOLKERLA 261

Db 197 COLNVLNDRKRTALIKAVQCEDECALMLLEHGTDPNI----- 234

QY 262 YFKGSHLLQAAREADVTRIKKHLSEMVNFKHPQTH-ETALHCAAAAPYPRKQICE 320

Db 235 -----PDEYNTTLHYAI---YNEKLMKAL 258

QY 321 LRKGANINEKTFELPLHVASEKAHNDVVVYVVEVYVVEVYVVEVYVVEVYVVEVYVVEVYV 380

Db 259 LLYGADIESKNKGLTPLLGVHKEQKQVVKFLIKKANKANLNDRYGRTALILAVCCGSA 318

QY 381 QTCRLLLSYGCDPNIIISLQFTALOMG-----NENVQOLLOE-----GISLGNSEAD-- 427

Db 319 SIVSLLLEQNIDSSQDLSGOTAREYAVSHHHVICOILSDYKEKQMLKTSSNSNPENV 378

QY 428 -----ROLLEAAKAGDVETVKLCTVOSVN-----CRD----- 455

Db 379 SRTNRKPRTHMVVEVDSMPAASSYKVPFGLRSKMGKWCRCFFCCRESGKSNVGTSGDHD 438

QY 456 ----- 455

Db 439 DSAMKTLRSKMGKWCRCFFCCRGSKSNVGASGDHDDSAMKTLRNMKGKWCRCFFCCPCCR 498

QY 456 -----IEGROSTPLHFAAGYNRVSVVVEYLLO-HGADVHAKD 490

Db 499 GSKSKVGAWGDYDDSAFMEPRYHVVRGDELDKLHRAAMWGKVPKDLIVMLRDTDVNKKD 558

QY 491 KGLVPLHNAACSYGHYVVEVYVVEVYVVEVYVVEVYVVEVYVVEVYVVEVYVVEVYVVEVYV 550

Db 559 KQKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCEDECALMLLEHGT 618

QY 551 DPTKKNRDGNTPLDLVKDGTDIQDLLRGDAALLDAAKGCLARVKKLSSPWNVNCRTQ 610

Db 619 DPNIPDEYGN----- 628

QY 611 GRHSTPLHAAAGYNNLEVAEYLLHOGADVNAQDKGLIPLHNAASYCHVDVAALLIKYNA 670

Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHLTPILLGVHKEQKQVVKFLIKKA 684

QY 671 CVNATDKWFTPLHAAAGKRTOLCALLAHGADPTLLNCHNKSALDIAPTPOLKERLA 730

Db 685 NLNADRYGRTALILAVCCGSASIVSLLLEQNIDVSSQDLSGOTAREYAVSSHVHI--- 741

QY 731 AMPSPALPCYKQVNLNGVRSPGATADALSSGSPSSLSAASLNDLSSGFSSELSSVWS 790

Db 742 -----COLLSYKQMLKISSE-----NSNPQDLKLTSEESQRFKGS----- 781

QY 791 SSGTEGASSLEKKEVPGVDFSIOTFVRNGLHLEHMLDIFEREQITLDVLVEMGHKELKEIG 850

Db 782 -----ENSOPEKMSQPEINKD-----GDRE-----VE-----EEMKKHE 811

QY 851 INAYGHRHKLKIGVERLISGOGLNPLYLTMTSGTILIDLSPPDKFQSVVEEMQSTV 910

Db 812 SNNVGLLENITNGV-TAGNDGNGLIPOKRSRT-----PENQOPDNSEETHRI 859

QY 911 REHRDGGHAGGIFNRYNLIKTKQVCNK-----KLWERYTHRRKVEYSEENHNH----- 957

Db 860 CE-----LVSDYKEKQMPKYSSNSNPEDLKTSEESQRFKGS----- 910

QY 958 ANERMLFHGSPFVNAITHKGFDE 980

Db 911 AIEEMKKHGS-----THVGPE 927

RESULT 9

US-09-895-793-378

; Sequence 378, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedrick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C2  
; CURRENT APPLICATION NUMBER: US/09/895,793  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 378  
; LENGTH: 1719  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-895-793-378

Query Match 6.6%; Score 379.5; DB 9; Length 1719;  
Best Local Similarity 20.9%; Pred. No. 5.6e-19;  
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;  
QY 143 GRKSTPLHLAGYNNRKYIVQLLQ-HGADVHAKDGLVPLHNACSYGHVEYVTELLVKG 201  
DB 137 GEDLDKLRHAAWGWKVPKRDILVMDRTDVKKKOKRKTALHLASANGSEVVKLLDRR 196  
QY 202 ACVNAMDLWQFTPLHFAASKNRVEVCSLLSYGADPTLLNCHNKSALDAPTOLKERLA 261  
DB 197 COLNVLDNKKRTALIKAVQCEDECALMLLEHGTDPNI----- 234  
QY 262 YEFKSHLSLQAAREADVTRIKKHLNLEWVNFKHPQTH-ETALHCAASPYPKRKQICELL 320  
DB 235 -----PDEYGNITLHYAI---YNEDKLMKAL 258  
QY 321 LRKGANINEKTEFTPLHFASEKAHNDVVEVVKHAKVNALDNLGOTSLHRAAYCGHL 380  
DB 259 LLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKANLDRYGRTRALILAVCCGSA 318  
QY 381 QTCRLLLSYCDPNIIISQGTALQMG-----NENVOQLQOE-----GISLGNSEAD-- 427  
DB 319 SIVSLLEQINIDVSSQDLSGQTAREYAVSSHVVICOLLSDYKEKOMLKISSENSPENV 378  
QY 428 -----ROLLEAAKAGDVETVKKLCTVQSVN-----CRD----- 455  
DB 379 SRTNRPRTHVVEVDSMPAASVVKPFGLRSMKMGKWCCHCFPCCRSGKSNVGTSGDHD 438  
QY 456 ----- 455  
DB 439 DSAMKTLRSKMGKWCCHCFPCCRSGKSNVGSAGSDHDSAMKTLRNKMGKWCCHCFPCR 498  
QY 456 -----IGRQSTPLHFAAGYNNRKYIVQLLQ-HGADVHAKD 490  
DB 499 GSKSKVGAWGDDYDGSFMEPRVHVGEDLDKLRHAAWGWKVPKRDILVMDRTDVKKKD 558  
QY 491 KGLVPLHNACSYGHVEYVTELLVKGAVVNVADLWFTPLHFAASKNRVEVCSLLSYGADPTLLNCHNKSALDAPTOLKERLA 550  
DB 559 KOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCEDECALMLLEHGT 618  
QY 551 DPTKKNRDNTPDLVKGDDTDIDQLLRGDAALLDAKGLARVKKLSPPDNVNCRDITQ 610  
DB 619 DNPIDPEYGN----- 628  
QY 611 GRHSTPLHLAGYNNRKYIVQLLQ-HGADVHAKDGLVPLHNACSYGHVEYVTELLVKG 670  
DB 629 -----TTLHYAIYNEDKLMKALLLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684  
QY 671 CVNATDKWAFTPLHFAASKNRVEVCSLLSYGADPTLLNCHNKSALDAPTOLKERLA 730  
DB 685 NLNALDRYGRTRALILAVCCGSAIYVSLLEQINIDVSSQDLSGQTAREYAVSSHVVIV----- 741  
QY 731 AMPPSALPSCYKPOVLNGVRPGATADALSSGSPSSLSAASSLDNLSGFSSELSSVVS 790  
DB 742 -----COLLSYDEKQKMLKISSE-----NSNPEQDLKLTSEESQRFKGS----- 781  
QY 791 SSGTEGASLEKKEVPGVDFSTQFVRNLGLEHMDIFEREQITLDLVLEMGHKELKEIG 850  
DB 782 -----ENSOPEKMSQPEINKD-----GDRE-----VE-----EEMKHE 811  
QY 851 INAYGHRHKLIGVERLISGQGLNPYTLNTSGSTILIDLSPDDKEFQSEVEEMQSTV 910

DB 812 SNNVGLLENLTNGV-TAGNGDNGLIPOQRKSR-----PENQOFPDNESEYHRI 859  
QY 911 REHRDGGHAGGIFNRYNLIKIKQVCNK-----KLWERYTHRRKKEVSEENHH----- 957  
DB 860 CE-----LVSDYKEKQMPKYSSENSPQDLKLTSEESQRLGSENGQPELENPM 910  
QY 958 ANERMLFHGSPFVNAILHKGDE 980  
DB 911 AIEEMKKHGS-----THVGFP 927  
RESULT 10  
US-09-895-814-378  
; Sequence 378, Application US/09895814  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Hepler, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C26  
; CURRENT APPLICATION NUMBER: US/09/895,814  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 990  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 378  
; LENGTH: 1719  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-895-814-378

Query Match 6.6%; Score 379.5; DB 9; Length 1719;  
Best Local Similarity 20.9%; Pred. No. 5.6e-19;  
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;  
QY 143 GRKSTPLHLAGYNNRKYIVQLLQ-HGADVHAKDGLVPLHNACSYGHVEYVTELLVKG 201  
DB 137 GEDLDKLRHAAWGWKVPKRDILVMDRTDVKKKOKRKTALHLASANGSEVVKLLDRR 196  
QY 202 ACVNAMDLWQFTPLHFAASKNRVEVCSLLSYGADPTLLNCHNKSALDAPTOLKERLA 261  
DB 197 COLNVLDNKKRTALIKAVQCEDECALMLLEHGTDPNI----- 234  
QY 262 YEFKSHLSLQAAREADVTRIKKHLNLEWVNFKHPQTH-ETALHCAASPYPKRKQICELL 320  
DB 235 -----PDEYGNITLHYAI---YNEDKLMKAL 258  
QY 321 LRKGANINEKTEFTPLHFASEKAHNDVVEVVKHAKVNALDNLGOTSLHRAAYCGHL 380  
DB 259 LLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKANLDRYGRTRALILAVCCGSA 318  
QY 381 QTCRLLLSYCDPNIIISQGTALQMG-----NENVOQLQOE-----GISLGNSEAD-- 427

```

Db 319 STVSLLEQNDIVSSQDLSGTAREYAVSSHVHHVICQLLSYKEQKMLKTSSENSENPNV 378
Qy 428 -----ROLLEAAKAGDVETVKLCTVQSVN-----CRD----- 455
Db 379 SRTRNKPRTHMVVEVDSMPAASSVKKPFLGRLSKMGKWCRCFCPCRESGKSNVGTSGDHD 438
Qy 456 ----- 455
Db 439 DSAMKTLRSKMGKWCRCFCPCRESGKSNVGSAGDHDDSAMKTLRNKMGKWCRCFCPCR 498
Qy 456 -----IEGRQSTPLHFAAGYNRVSVVEYLQ-HGADVHAKD 490
Db 499 GSGKSVGAWGDYDSDAFMEPRYHVRGEDLDKLHRAAWGKVPKDLIVMLRDTDVNKKD 558
Qy 491 KGGVPLHNACSYGHYEAELLVKGAVNVADLWKFTPLHEAAAKGYEICKLLLOHGA 550
Db 559 KOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVOCQDECALMLEHGT 618
Qy 551 DPTKKNRGDNTPLDLVKOGDTDIQLLRGDAALLDAAKGGLARVKKLSPPDNVNCRTQ 610
Db 619 DPNIPDEYGN----- 628
Qy 611 GRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNA 670
Db 629 ----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684
Qy 671 CYNATDKWAFTPLHEAAKGRQTQCALHLAHGADPTLNQEGOTPLDLVSDADDVSALLTA 730
Db 685 NLNADRYGRTALILAVCCGSASIVSLLEQNDIVSSQDLSGTAREYAVSSHVHHVI--- 741
Qy 731 AMPPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSEFSELSSVVS 790
Db 742 ----COLLSDYKEKQMLKTSSE-----NSNPEODLKLITSEESORFKGS----- 781
Qy 791 SSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
Db 782 ----ENSQPEKMSOEPEINKD-----GDRE-----VE-----EEMKKHE 811
Qy 851 INAYGHRHKLKIGVERLISGOQGLNPLYLTNTSGSGTILIDSPDDKEFQSVSEEMQSTV 910
Db 812 SNNVGLLENLTNGV-TAGDNGNLIPQKRSRT-----PENQOPFDNESEYHRI 859
Qy 911 REHRDGGHAGGIFNRYNLIKIOKVCNK-----KLWERYTHRRKEVSEENHHN----- 957
Db 860 CE-----LVSDYKEKQMPKYSSENSENPEODLKLITSEESORLEGGSENGOPELENFM 910
Qy 958 ANERMLFHGSPFFVNAIHHKGFDE 980
Db 911 AIBEMKKHGS-----THVGFE 927

```

## RESULT 11

US-09-759-143-378  
 ; Sequence 378, Application US/09759143  
 ; Patent No. US2002022248A1

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Harlocker, Jennifer L.  
 APPLICANT: Harlocker, Susan L.  
 APPLICANT: Jiang, Yuqi  
 APPLICANT: Henderson, Robert A.  
 APPLICANT: Kalos, Michael D.  
 APPLICANT: Fanger, Gary R.  
 APPLICANT: Retter, Marc W.  
 APPLICANT: Stolk, John H.  
 APPLICANT: Day, Craig H.  
 APPLICANT: Vedvick, Thomas S.  
 APPLICANT: Carter, Darrick  
 APPLICANT: Li, Samuel  
 APPLICANT: Wang, Aijun  
 APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Hepler, William  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 FILE REFERENCE: 210121.427C23  
 CURRENT APPLICATION NUMBER: US/09/759,143  
 CURRENT FILING DATE: 2001-01-12  
 NUMBER OF SEQ ID NOS: 934  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 378  
 LENGTH: 1719  
 TYPE: PRT  
 ORGANISM: Homo sapien  
 US-09-759-143-378

Query Match 6.6%; Score 379.5; DB 10; Length 1719;

Best Local Similarity 20.9%; Pred. No. 5.6e-19;

Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

Qy 143 GRKSTPLHLAAGYNRVKIVQLLIQ-HGADVHAKDGLVPLHNACSYGHYEVTELLVKHG 201

Db 137 GEDLDKLHRAAWGKVPKDLIVMLRDTDVNKKDKQKRTALHLASANGSEVVKLLDR 196

Qy 202 ACYNAMDLOFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSALDLAPTQLKERLA 261

Db 197 CQLNVLDNKKRTALIKAVOCQDECALMLEHGTDPNI----- 234

Qy 262 YEFKHSILQAAAREADVTRIKKHLSEWVNFKHPQTH-ETALHCAAAASPYPKRKQICELL 320

Db 235 -----PDEYGNITLHYAI---YNEDKLMAKAL 258

Qy 321 LRKGANINEKTEFLTPLHVASEKAHNDVVEVVKHAKVNALDNLGOTSILHRAAYCGLH 380

Db 259 LLYGADTESKNKHGLTPLLGVHEQKQVVKFLIKKANLALNADRYGRTALILAVCCGSA 318

Qy 381 QTCRLLSYCGDPNIIISOGETALQMG-----NENVOQLLOP-----GISLGNSEAD-- 427

Db 319 SIVSLLEQNDIVSSQDLSGTAREYAVSSHVHHVICQLLSYKEQKMLKTSSENSENPNV 378

Qy 428 -----ROLLEAAKAGDVETVKLCTVQSVN-----CRD----- 455

Db 379 SRTRNKPRTHMVVEVDSMPAASSVKKPFLGRLSKMGKWCRCFCPCRESGKSNVGTSGDHD 438

Qy 456 ----- 455

Db 439 DSAMKTLRSKMGKWCRCFCPCRESGKSNVGSAGDHDDSAMKTLRNKMGKWCRCFCPCR 498

Qy 456 -----IEGRQSTPLHFAAGYNRVSVVEYLQ-HGADVHAKD 490

Db 499 GSGKSVGAWGDYDSDAFMEPRYHVRGEDLDKLHRAAWGKVPKDLIVMLRDTDVNKKD 558

Qy 491 KGGVPLHNACSYGHYEAELLVKGAVNVADLWKFTPLHEAAAKGYEICKLLLOHGA 550

Db 559 KOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVOCQDECALMLEHGT 618

Qy 551 DPTKKNRGDNTPLDLVKOGDTDIQLLRGDAALLDAAKGGLARVKKLSPPDNVNCRTQ 610

Db 619 DPNIPDEYGN----- 628

Qy 611 GRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNA 670

Db 629 ----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684

Qy 671 CYNATDKWAFTPLHEAAKGRQTQCALHLAHGADPTLNQEGOTPLDLVSDADDVSALLTA 730

Db 685 NLNADRYGRTALILAVCCGSASIVSLLEQNDIVSSQDLSGTAREYAVSSHVHHVI--- 741

Qy 731 AMPPSALPCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSEFSELSSVVS 790

Db 742 ----COLLSDYKEKQMLKTSSE-----NSNPEODLKLITSEESORFKGS----- 781

Qy 791 SSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850

Db 782 ----ENSQPEKMSOEPEINKD-----GDRE-----VE-----EEMKKHE 811

QY 851 INAYGHRHKLKIGVERLISGQGLNPLYLTNTSGSTILIDLSPDDKEFQSEVEEMOSTV 910  
Db 812 SNNVGLLENLTNGV-TAGNGDGLIPQKRSRT-----PENQOFPDNESEYHRI 859  
QY 911 REHRDGGHAGGIFNRYNLIKTKQVCKN-----KLWERYTHRRKEVSEENHNH----- 957  
Db 860 CE-----LVSDYKEKQMPKYSSNSNPEDQLKLTSEESQRLGSENGQPELENFM 910  
QY 958 ANERMLFHGSPFVNAIHKGFDE 980  
Db 911 AIEEMKKHGS-----THVGPE 927

RESULT 12  
US-09-780-669-378  
; Sequence 378, Application US/09780669  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C24  
; CURRENT APPLICATION NUMBER: US/09/780,669  
; CURRENT FILING DATE: 2001-02-09  
; NUMBER OF SEQ ID NOS: 943  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 378  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-780-669-378

Query Match 6.6%; Score 379.5; DB 10; Length 1719;  
Best Local Similarity 20.9%; Pred. No. 5.6e-19;  
Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

QY 143 GRKSTPLHLAAGNRYKIVQLLLO-HGADVHAKDGLVPLHNACSYGHYEVTELLVKHG 201  
Db 137 GEDLDKLHRAAMWGKVPKRDLIIVMLRDTDVNKKDKOKRTALHLASANGNSEVVKLLDDR 196  
QY 202 ACVNMDLMOFTPLHFAASKNRVEVCSLLLSYGADPTLLNCHNKSALDIATPPQLKERLA 261  
Db 197 COLNVLDNKKRTALIKAVQCOQDECALMLEHGTDPNI----- 234  
QY 262 YEFKSHLSLQAAREADVTRIKKHSLEMVNFVKHPOTH-ETALHCAAASPYPKRQICELL 320  
Db 235 -----PDEYGNITLHVAI-----YNEKLMKAKAL 258  
QY 321 LRKGANINEKTEFTPLHFASEKANDVDVEVVKHEAKVNALDNLGOTSILHRAAYCGHL 380  
Db 259 LLYGADIESKNKHLGTPLLLVGHEQKQVVKPLIKKANLANALORYGRTALILAVCCGSA 318  
QY 381 QTCRLLLSYGCDPNIISLQGTALQMG-----NENVQQLQOE-----GISLGNSEAD-- 427

Db 319 SIYVLLLEQNIDYSSODLSGQTAREYAVSSHVHHVICQLLSDYKEKQMLKISSNSPENV 378  
QY 428 -----POLLEAAKAGDVETVKKICTVOSVN-----CRD----- 455  
Db 379 SRTNKPRTHMVVEVDSMPAASVKKPFGRLSRKMGKWCRCPCPCRESKSNVGTSGDHD 438  
QY 456 ----- 455  
Db 439 DSAMKTLRSKMGKWCRCPCPCRCGSKSNVGASGDHDDDSAMKTLRNMKGKWCCHCPCCR 498  
QY 456 -----TEGROSTPLHFAAGYNRVSVVEYLQ-HGADVHAKD 490  
Db 499 GSGSKVGAWGDDVSFAFMPEYHVRGEDLDKLHRAAMWGKVPKRDLIIVMLRDTDVNKKD 558  
QY 491 KGLVPLHNACSYGHYEAELLVKHGAVNVADLMKFTPLHEAAAKGKYEICKLLOHCA 550  
Db 559 KOKRTALHLASANGNSEVVKLLLDRCQLNVLDNKKRTALIKAVQCOQDECALMLEHGT 618  
QY 551 DPTKKNRDGNTPLDIVKGDGTDIQLLRGDAALLDAKKGCLARVKKLSSPDVNCRDQTQ 610  
Db 619 DPNIPDEYN----- 628  
QY 611 GRHSTPLHLAAGYNMLEVAEYLLOHCAOVNAQDKGLIPLHNAASYGHVDVAALLIKYNA 670  
Db 629 -----TTLHYAIYNEKLMKAKALLYGADIESKNKHGLTPLLGVHEQKQVVKFLIKKKA 684  
QY 671 CVNATDKWAFPLHFAAOKGRTQOLCALLLAHGAADPTLKNQEGOTPLDLVSADDSALLTA 730  
Db 685 NLNALDRYGRITALILAVCCGSASIVSLLLEQNIYSSQDLSQOTAREYAVSSHVHI--- 741  
QY 731 AMPPSALPCYKPOVLINGVRSPGATADALSSGSPSPSSLSAASSLDNLGSEFSELSSVVS 790  
Db 742 -----COLLSDYKEKQMLKTSSE-----NSNPEODLKLTSEESQRFKGS----- 781  
QY 791 SSGTEGASSLEKKEVPVGVDFSITQFVRLGLEHLMDIFEREQITLDVLVEMGHKELKEIG 850  
Db 782 -----ENSQPEKMSQPEINKD-----GDRE-----VE-----EEMKKHE 811  
QY 851 INAYGHRHKLKIGVERLISGQGLNPLYLTNTSGSTILIDLSPDDKEFQSEVEEMOSTV 910  
Db 812 SNNVGLLENLTNGV-TAGNGDGLIPQKRSRT-----PENQOFPDNESEYHRI 859  
QY 911 REHRDGGHAGGIFNRYNLIKTKQVCKN-----KLWERYTHRRKEVSEENHNH----- 957  
Db 860 CE-----LVSDYKEKQMPKYSSNSNPEDQLKLTSEESQRLGSENGQPELENFM 910  
QY 958 ANERMLFHGSPFVNAIHKGFDE 980  
Db 911 AIEEMKKHGS-----THVGPE 927

## RESULT 13

US-09-822-827-378  
; Sequence 378, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.534C1  
; CURRENT APPLICATION NUMBER: US/09/822,827  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 378  
; LENGTH: 1719  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-822-827-378

Query Match 6.6%; Score 379.5; DB 10; Length 1719;  
Best Local Similarity 20.9%; Pred. No. 5.6e-19;

Matches 205; Conservative 106; Mismatches 335; Indels 337; Gaps 23;

QY 143 GRKSTPLHLAAGYNNKIVQLLLO-HGADVHAKDKGLDVLPLHNACSYGHVEVTELLVKG 201  
DB 137 GEDLKLHRAAWGKVPKRDILVLMRLDVTNKKDKOKRTALHLASANGSEVVKLLDRR 196  
QY 202 ACVNAMDLMQFTPLHEAAKSKRVVCSLLLSYGADPTLLNCHNKSALDAPTOLKERLA 261  
DB 197 CQLNVLNKKRTALIKAVQCCDECALMLLEHGTDPNT----- 234  
QY 262 YEFKGSLLQAAAREADVTRIKKLSLEMVNFKHPOTH-ETALHCAAAAPYPRKKOICELL 320  
DB 235 LRRGAMINEKTEKFTPLHVASEKAHNDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHL 258  
QY 321 LRRGAMINEKTEKFTPLHVASEKAHNDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHL 380  
DB 259 LLYGADIESKNKHGTLPLLLGVHEQKQVYKFLIKKANLANALDRYGRALILAVCCGSA 318  
QY 381 QTCRLLSYGCDPNIIISLOGFTALONG-----NENVOQLLOE-----GISLGNSEAD-- 427  
DB 319 SIVSLLLEQNIIDVSSQDLSGQTAREYAVSSHHVHICOLLSDYKEKOMLKISSENSPENV 378  
QY 428 -----RQLEAAKAGDVETVKKLCITVQSVN-----CRD----- 455  
DB 379 SRTNKPRTMVMVEVDSPMAASVKKPFLRSKMGKWCRCFPCCRESKSNVGTSGDHD 438  
QY 456 ----- 455  
DB 439 DSAMKTLRSKMGKWCRCFPCCRGSKSNVAGSGDHDSDAMKTLRNKMGKWCRCFPCCR 498  
QY 456 ----- 498  
DB 499 GSKSKVAGWDYDYSAFMEPRYHVRGDELDKLHRAAWGKVPKRDILVLMRLDVTNKKD 558  
QY 491 KGGVLPLHNACSYGHVEAELLVKGAVNVADLMKFTPLHEAAKSKRGYKICOLLQHA 550  
DB 559 KQKRTALHLASANGSEVVKLLDRRCQLNVLNKKRTALIKAVQCCDECALMLLEHGT 618  
QY 551 DPTKKNRDNWTPDLVKGDTDIDQLLRGDAALLDAKKGCLARVKKLSSPDNVNCRDQTQ 610  
DB 619 DPNIPDEYGN----- 628  
QY 611 GRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNA 670  
DB 629 -----TTLHYAIYNEDKLMALIKYLGADIESKNKHGTLPLLLGVHEQKQVVKFLIKKA 684  
QY 671 CVNATDKWFTPLHEAAKGRTOLCALLLAHAGADPTLKNQEGOTPLDLVSADVSALLTA 730  
DB 685 NLNALDRYGRALILAVCCGSAISVLLLEQNIIDVSSQDLSGQTAREYAVSSHHHVI--- 741  
QY 731 AMPSALPSCYKQVNLGVNRPNGATADALSSGSPSPSSLSAASLDNLGSGFSSELSSVVS 790  
DB 742 -----CQLSDYKEKOMLKISSE-----NSNPEQDLKLTSEESORFKGS----- 781  
QY 791 SSGTGASSLEKKBVGVDFESITQFVRNLGLEHLMIDIFEREQITLDVLVEMGHKELKEIG 850  
DB 782 -----ENSOPEKMSQPEIND-----GDRE-----VE-----EEMKKHE 811  
QY 851 INAYGHRHKLKIGVERLLISGOGLNPLYLTNTSSGTLILDLSPDDKEFQSVSEEMQSTV 910  
DB 812 SNNVGLLENLTNGV-TAGNGDNGLIPOKRSKT-----PENQOFPDNESEYHRI 859  
QY 911 REHRDGGHAGGIFRNYITLTKQVCNK-----KLWERYTHRRKKEVSEENHNH----- 957  
DB 860 CE-----LVSDYKEKOMPKYSSSENSNPQDLKLTSEESORLESGSENGQPELENFM 910  
QY 958 ANERMLFHPGSPFVNAIHKGFDE 980  
DB 911 ATEEMKKHGS-----THVGPE 927

Query Match 6.4%; Score 368.5; DB 10; Length 740;  
Best Local Similarity 22.0%; Pred. No. 9.8e-19;  
Matches 167; Conservative 84; Mismatches 246; Indels 261; Gaps 21;

QY 3 GRKDWVEVLLQ---NGASVOARDGGLIPLHNACSFHAEVNVNLLLRHGAD-----PN 52  
DB 138 GOKOTVSLHQCSHKGVPFEGEGDGLSPEDH-----FSELPOVDITLFDNDVDDDEQSQSP 193  
QY 53 ARDNMNYTPLHEAA-----TKGKIDVC-----IVLLQHGAEPTI 86  
DB 194 SAEQIDFVVPQPLSSPQCNSFSSDLGSGNCTNSLELQKVSNGQIVGQPOIATITDGOGLLV 253  
QY 87 RNTDG-----RTALDLADPSAKAVLTGEYKKBELLESABSGNEEKMMALLTPLNVNCH 139  
DB 254 QEPDGLMVATPAQTUTDLDLIAAVST-----RVPTGNSSSSQTECLTPESCQOT 305  
QY 140 ASD-GRKSTPLHLAAGYNNRVKIVQLLQHGADVHAKDKGLVPLHNACSYGHVEVTELLV 198  
DB 306 TSNVASQSNP-----PVYPSVDI-----DAHTESNHD-TALTACAGGHEELVSLVI 351  
QY 199 KHGACVNAMDLMQFTPLHEAAKSKRVVCSLLLSYGADPTLLNCHNKSALDAPTOLKE 258  
DB 352 ARDAKIEHRDKKGFPTLILATAGHGVVVEILLDDKGD----- 389  
QY 259 RLAYEFKGSLLQAAAREADVTRIKKLSLEMVNFKHPOTHETALHCAAAAPYPRKQICE 318  
DB 390 -----TEAQSE-----RTKDTPLSLACSG-----GRQEVVD 415  
QY 319 LLLRKGANINEKTEKFTPLHVASEKAHNDVVEVVKHEAKVNAL--DNLGOTSILHRAAY 376  
DB 416 LLLARGANKEHRNVSDYTPLSLAASGGYVNIITIKILNAGAEINSTRGSKLIGISPLMLAAM 475  
QY 377 CGHLQTCRLLSYGCDPNIIISLOGFTALONGNENVOQLLOEGISLGNSEADQLLEAARA 436  
DB 476 NGHVPVAVKLLDNGSDIN----- 493  
QY 437 GUVETVKKLTCTVQSVNCRDIEGROSTPLHFAAGYNNRVSVVEYLLQHGADVHAKDKGLVP 496  
DB 494 -----AQIETNRTALTALACFOGRAEVVSLLLDRKANVEHRAKTGLTP 536  
QY 497 LHNACSYGHVEAELLVKGAVNVADL--WKFTPLHEAAKSKRGYKICOLLQHGADPTK 554  
DB 537 LMEASGGYAEVGRVLLDKGADVNAAPPVSSRDALTIAADKGYKFCCELLTHRGAHIDV 596  
QY 555 KNRDGNTPDLVKGDTDIDQLLRGDAALLDAKKGCLARVKKLSSPDNVNCRDQTGRHS 614  
DB 597 RNKKGNTPLNASG----- 611  
QY 615 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK----- 667

QY 143 GRKSTPLHLAAGYNNKIVQLLLO-HGADVHAKDKGLDVLPLHNACSYGHVEVTELLVKG 201  
DB 137 GEDLKLHRAAWGKVPKRDILVLMRLDVTNKKDKOKRTALHLASANGSEVVKLLDRR 196  
QY 202 ACVNAMDLMQFTPLHEAAKSKRVVCSLLLSYGADPTLLNCHNKSALDAPTOLKERLA 261  
DB 197 CQLNVLNKKRTALIKAVQCCDECALMLLEHGTDPNT----- 234  
QY 262 YEFKGSLLQAAAREADVTRIKKLSLEMVNFKHPOTH-ETALHCAAAAPYPRKKOICELL 320  
DB 235 LRRGAMINEKTEKFTPLHVASEKAHNDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHL 258  
QY 321 LRRGAMINEKTEKFTPLHVASEKAHNDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHL 380  
DB 259 LLYGADIESKNKHGTLPLLLGVHEQKQVYKFLIKKANLANALDRYGRALILAVCCGSA 318  
QY 381 QTCRLLSYGCDPNIIISLOGFTALONG-----NENVOQLLOE-----GISLGNSEAD-- 427  
DB 319 SIVSLLLEQNIIDVSSQDLSGQTAREYAVSSHHVHICOLLSDYKEKOMLKISSENSPENV 378  
QY 428 -----RQLEAAKAGDVETVKKLCITVQSVN-----CRD----- 455  
DB 379 SRTNKPRTMVMVEVDSPMAASVKKPFLRSKMGKWCRCFPCCRESKSNVGTSGDHD 438  
QY 456 ----- 455  
DB 439 DSAMKTLRSKMGKWCRCFPCCRGSKSNVAGSGDHDSDAMKTLRNKMGKWCRCFPCCR 498  
QY 456 ----- 498  
DB 499 GSKSKVAGWDYDYSAFMEPRYHVRGDELDKLHRAAWGKVPKRDILVLMRLDVTNKKD 558  
QY 491 KGGVLPLHNACSYGHVEAELLVKGAVNVADLMKFTPLHEAAKSKRGYKICOLLQHA 550  
DB 559 KQKRTALHLASANGSEVVKLLDRRCQLNVLNKKRTALIKAVQCCDECALMLLEHGT 618  
QY 551 DPTKKNRDNWTPDLVKGDTDIDQLLRGDAALLDAKKGCLARVKKLSSPDNVNCRDQTQ 610  
DB 619 DPNIPDEYGN----- 628  
QY 611 GRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNA 670  
DB 629 -----TTLHYAIYNEDKLMALIKYLGADIESKNKHGTLPLLLGVHEQKQVVKFLIKKA 684  
QY 671 CVNATDKWFTPLHEAAKGRTOLCALLLAHAGADPTLKNQEGOTPLDLVSADVSALLTA 730  
DB 685 NLNALDRYGRALILAVCCGSAISVLLLEQNIIDVSSQDLSGQTAREYAVSSHHHVI--- 741  
QY 731 AMPSALPSCYKQVNLGVNRPNGATADALSSGSPSPSSLSAASLDNLGSGFSSELSSVVS 790  
DB 742 -----CQLSDYKEKOMLKISSE-----NSNPEQDLKLTSEESORFKGS----- 781  
QY 791 SSGTGASSLEKKBVGVDFESITQFVRNLGLEHLMIDIFEREQITLDVLVEMGHKELKEIG 850  
DB 782 -----ENSOPEKMSQPEIND-----GDRE-----VE-----EEMKKHE 811  
QY 851 INAYGHRHKLKIGVERLLISGOGLNPLYLTNTSSGTLILDLSPDDKEFQSVSEEMQSTV 910  
DB 812 SNNVGLLENLTNGV-TAGNGDNGLIPOKRSKT-----PENQOFPDNESEYHRI 859  
QY 911 REHRDGGHAGGIFRNYITLTKQVCNK-----KLWERYTHRRKKEVSEENHNH----- 957  
DB 860 CE-----LVSDYKEKOMPKYSSSENSNPQDLKLTSEESORLESGSENGQPELENFM 910  
QY 958 ANERMLFHPGSPFVNAIHKGFDE 980  
DB 911 ATEEMKKHGS-----THVGPE 927

RESULT 14  
US-09-835-788A-12

Search completed: February 12, 2003, 07:52:59  
Job time : 27.5641 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:07:14 : Search time 21.6239 Seconds  
(without alignments)  
4890.316 Million cell updates/sec

Title: US-09-843-159b-3

Perfect score: 5769

Sequence: 1 GGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPGMVDG 1100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	877.5	15.2	1765	2 T42714	ankyrin 3, splice
2	877.5	15.2	1961	2 T42716	ankyrin 3, splice
3	875.5	15.2	4377	2 A55575	ankyrin 3, long sp
4	873.5	15.1	1943	2 T42713	ankyrin 3, splice
5	871.5	15.1	1940	2 T42715	ankyrin 3, splice
6	860.5	14.9	3924	2 S37431	ankyrin 2, neurona
7	844	14.6	1856	2 B35049	ankyrin 1, erythro
8	844	14.6	1880	2 A35049	ankyrin 1, erythro
9	842	14.6	1881	1 SJHUK	ankyrin 1, erythro
10	839.5	14.6	1848	2 S37771	ankyrin, erythrocy
11	837.5	14.5	1862	2 T49502	ankyrin - mouse
12	834	14.5	1549	2 T13940	ankyrin - fruit fl
13	814.5	14.1	2039	2 T15347	ankyrin-related un
14	628	10.9	791	2 T42691	hypothetical prote
15	545.5	9.5	1411	2 S30355	alpha-latroinsecto
16	514	8.9	934	1 H71274	probable ankyrin -
17	496.5	8.6	1062	2 T30255	inversin - mouse
18	491.5	8.5	1062	2 T14151	Inv protein - mous
19	464.5	8.1	1401	2 S11527	alpha-latrotoxin p
20	456.5	7.9	1188	2 T19552	hypothetical prote
21	451	7.8	1058	2 D82654	ankyrin-like prote
22	431.5	7.5	426	2 AE2149	hypothetical prote
23	426	7.4	991	2 T23412	hypothetical prote
24	405.5	7.0	1184	2 T00253	gene Ankhzn protei
25	388	6.7	397	2 T46445	hypothetical prote
26	372.5	6.5	2584	2 T24158	hypothetical prote
27	372.5	6.5	2806	2 T24157	hypothetical prote
28	363.5	6.3	1031	2 T43458	hypothetical prote
29	350	6.1	557	2 T46507	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

T42714

ankyrin 3, splice form 2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000

C:Accession: T42714

R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge the repeat domain.

A:Reference number: Z22237; MUID:95340633; PMID:7615634

A:Accession: T42714

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1765 <PET>

A:Cross-references: EMBL:L40632; NID:G710548; PID:G710551; PIDN:AAB01605.1

C:Experimental source: strain C57BL/6J; Kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 1587/1

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 15.2%; Score 877.5; DB 2; Length 1765;  
Best Local Similarity 30.9%; Pred. No. 1.8e-46;  
Matches 277; Conservative 125; Mismatches 377; Indels 117; Gaps 22;

QY 3 GRKDVVEYLLQNGASVQARDGGLPLHNACSFGEAEVYNLLLRHGADPNARDNNVTPL 62

Db 68 GHVEVVSSELLQREANVDAATKKGNTALHSLAQAEVVKVLVTNGANVNAQSQNGFTPL 127

QY 63 HEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDIA-----DPSAKAVLTGEYK---KDE 114

Db 128 YMAAGENHLEVVRFLDNGASQSLATEDGFTPLAVLQQHQDVVSVLLENDTIGKVKRLP 187

QY 115 LLESARSNEEKMMALLPLPNVNCASDGRKSTPLHLAAGYNRYKIVOLLQHGADVHAK 174

Db 188 ALHIAARKDDTKAAALLQNDTNAVDESQSGFTPLHIAAHYGNINVTALLNRAANVDF 247

QY 175 DKGLDVLPLHNACSYGHEVTELLVVKHGACVNMADLWQFTPLHEAASKNRVEYCSLLSYG 234

Db 248 ARNDITPLHVASKRGNNANVKKLLDRGAKIDAKTRDGLTPLHCGARSHEQVEMVLLDRS 307

QY 235 ADPTLLNCHNSAIDIAPTPQLKRLAYEFKGSHLSLQARADVTIRIKKHLISLEWNVFKH 294

Db 308 A-PILSKTKN-----GLSPLHMTQGD-----HLNCVOLLQ 339

QY 295 -----PQTHETALHCAASPYPKRKOICELLRLKGANINEKTEFTPLHVASEKAHN 347

Db 340 NVFVDVDTNDYLTALHVAHCHGYK---VAKVLLDKKASPNKALNGFTPLHIACKKNRI 396

hypothetical prote  
death-associated p  
hypothetical prote  
hypothetical prote  
hypothetical prote  
myosin heavy chain  
2-5A-dependent RNA  
ankyrin related pr  
hypothetical prote  
probable ankyrin  
transcription fact  
myosin-light-chain  
sex-determining pr  
hypothetical prote  
probable membrane  
related to 26s pro

QY 348 DVVEVVVKEAKYNALDNLGQTSIHHAAAGYGHLOTCRLLSYGCDDPNIIISLOGFTALOMG 407  
 Db 397 RVMEILLKHGASIOAVTESGLTPIHVAAFMGHVNIVSQLMHGASPNNTVNRGETALHMA 456  
 QY 408 -----NENVOOLLOEGISLGNSEADQ--LLEAAKAGDVETVKKLCTV-OSVNCRDIEGR 459  
 Db 457 ARSGQAEVVRVLYVDGGAQVEAKAKDDQTPLHISARLGKADIVQOOLLQOGASPNNAATTS- 515  
 QY 460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVPLHNAACSYGHYVAELLVKHCAYV 519  
 Db 516 -YTPLHAAAREGHEDVAFLLDHGASISITTKGFTPLHVAAYKGLVAVASLLLOKSASP 574  
 QY 520 NVADLWKFTPLHPEAAKGYEICKLLOHGADPTKKNRDNTPDL-LVKDGDITDIDL- 577  
 Db 575 DAAGKSGLTPLHVAAYHYNQKVALLLDQOGASPHAAAKNGYTPHIAAKKNQMDIATSL 634  
 QY 578 -----RDAALLDAAKKGCLARYKLLSPD--NNVCRDTQGRHSTPLHLAAGYNN 625  
 Db 635 EYGADANAVTROGIAVSHLAAQEGHVDVMSLLSRANVNLNKS- -LTPHLAAQEDR 692  
 QY 626 LEVAEYLLQHGADVNAODKGGVPLHNAASYGHDVVAALLIKYNACVNATDKWAFPLHE 685  
 Db 693 VNVAEVLVNOGAHVDAQTKMGYTPLVHGYGNKIVNVLQHSKVNAKTKNGYALHQ 752  
 QY 686 AAKGRTQLCALLAHGADPTLKNQEGQTPDLVS-----ADDYSALLITAMPSPALP 738  
 Db 753 AAQOGHTHIINVLQNNASPNELTVNGNTALATARRLGYSIVVDTLKVVTETIMTTTIT 812  
 QY 739 SCYK---POVLNCGVRSPGATADALS- -GSPSSLSAASLDNLS-GSFSLSVSVSSGT 794  
 Db 813 EXHKMNPETMEVLD-----MSDDEVKASAPKLSDEY-----ISDGE 853  
 QY 795 EGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850  
 Db 854 EGDKCTWFK-IPKQVEVLK-----SEDAITGDTDKYLGPDQLKELG 894  
  
 RESULT 2  
 T42716  
 ankyrin 3, splice form 4 - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
 C:Accession: T42716  
 R:Peters, L.D.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, I.  
 J. Cell Biol. 130, 313-330, 1995  
 A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
 the repeat domain.  
 A:Reference number: 222237; MUID:95340633; PMID:7615634  
 A:Accession: T42716  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1961 <PET>  
 A:Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1  
 A:Experimental source: strain C57BL/6J; kidney  
 C:Genetics:  
 A:Gene: Ank3  
 A:Map position: 10  
 C:Superfamily: ankyrin; ankyrin repeat homology  
 C:Keywords: alternative splicing  
  
 Query Match 15.2%; Score 877.5; DB 2; Length 1961;  
 Best Local Similarity 30.9%; Pred. No. 2.le-46;  
 Matches 277; Conservative 125; Mismatches 377; Indels 117; Gaps 22;  
  
 QY 3 GRKDVVEYLLONGASVQARDGGLIPLHNAACSFHAEVYVNLRLRHGADPNARDNNWYTPL 62  
 Db 68 GIVEVVELLQREANVDRAATKKNLNTALHTASLAGOAEVYKVLVTGANNVNAQOSNGFTPL 127  
 QY 63 HPAALIKGIDVCIVLLOHGAETPIRNTQRTALDLA-----DPSAKAVITGEYK---KDE 114  
 Db 128 YMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQOCHQDVVYSLLENKTKGVRLP 187  
 QY 115 LLESARSNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLLOHGADVHAK 174

Db 188 ALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVAITLLNRAAAVDFT 247  
 QY 175 DKGDVPLHNAACSYGHYVTELLVKGACVYNDLWQFTPLHHEAASKNRVYEGVSLLSYG 234  
 Db 248 ARNDITPLHNAACSYGHYVTELLVKGACVYNDLWQFTPLHHEAASKNRVYEGVSLLSYG 307  
 QY 235 ADPTLLNCHNKSAIDLAAPTLPOLKERLAYEFKSHLSLQAAAREADVTRIKKHLSLEWVNFKH 294  
 Db 308 A-PILSKTKN-----GLSPLHMAATQGD-----HLNVCVQLLQ 339  
 QY 295 -----PQTHETALHCAASPYKRRQICELLRLRGANINEKTEKFTPLHVAASEKAHN 347  
 Db 340 NVPVDVDDYNDLTAALHVAACHGHYK---VAKVLLDKKASPNALNGFTPLHIAACKNRI 396  
 QY 348 DVVEVVVKEAKYNALDNLGQTSIHHAAAGYGHLOTCRLLSYGCDDPNIIISLOGFTALOMG 407  
 Db 397 RVMEILLKHGASIOAVTESGLTPIHVAAFMGHVNIVSQLMHGASPNNTVNRGETALHMA 456  
 QY 408 -----NENVOOLLOEGISLGNSEADQ--LLEAAKAGDVETVKKLCTV-OSVNCRDIEGR 459  
 Db 457 ARSGQAEVVRVLYVDGGAQVEAKAKDDQTPLHISARLGKADIVQOOLLQOGASPNNAATTS- 515  
 QY 460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVPLHNAACSYGHYVAELLVKHCAYV 519  
 Db 516 -YTPLHAAAREGHEDVAFLLDHGASISITTKGFTPLHVAAYKGLVAVASLLLOKSASP 574  
 QY 520 NVADLWKFTPLHPEAAKGYEICKLLOHGADPTKKNRDNTPDL-LVKDGDITDIDL- 577  
 Db 575 DAAGKSGLTPLHVAAYHYNQKVALLLDQOGASPHAAAKNGYTPHIAAKKNQMDIATSL 634  
 QY 578 -----RDAALLDAAKKGCLARYKLLSPD--NNVCRDTQGRHSTPLHLAAGYNN 625  
 Db 635 EYGADANAVTROGIAVSHLAAQEGHVDVMSLLSRANVNLNKS- -LTPHLAAQEDR 692  
 QY 626 LEVAEYLLQHGADVNAODKGGVPLHNAASYGHDVVAALLIKYNACVNATDKWAFPLHE 685  
 Db 693 VNVAEVLVNOGAHVDAQTKMGYTPLVHGYGNKIVNVLQHSKVNAKTKNGYALHQ 752  
 QY 686 AAKGRTQLCALLAHGADPTLKNQEGQTPDLVS-----ADDYSALLITAMPSPALP 738  
 Db 753 AAQOGHTHIINVLQNNASPNELTVNGNTALATARRLGYSIVVDTLKVVTETIMTTTIT 812  
 QY 739 SCYK---POVLNCGVRSPGATADALS- -GSPSSLSAASLDNLS-GSFSLSVSVSSGT 794  
 Db 813 EXHKMNPETMEVLD-----MSDDEVKASAPKLSDEY-----ISDGE 853  
 QY 795 EGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850  
 Db 854 EGDKCTWFK-IPKQVEVLK-----SEDAITGDTDKYLGPDQLKELG 894  
  
 RESULT 3  
 A55575  
 ankyrin 3, long splice form - human  
 N:Alternate names: ankyrin G  
 C:Species: Homo sapiens (man)  
 C>Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 20-Sep-1999  
 C:Accession: A55575  
 R:Kordeli, E.; Lambert, S.; Bennett, V.  
 J. Biol. Chem. 270, 2352-2359, 1995  
 A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the  
 A:Reference number: A55575; MUID:95138209; PMID:7836469  
 A:Accession: A55575  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-4377 <KOR>  
 A:Cross-references: GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025  
 C:Genetics:  
 A:Gene: GDB:ANK3  
 A:Cross-references: GDB:424503; OMIM:600465  
 A:Map position: 10q21-10q21  
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol



C:Keywords: alternative splicing; peripheral membrane protein

	Query Match	15.2%	Score 875.5;	Db 2;	Length 4377;
	Best Local Similarity	30.5%	Pred. No. le=45;		
	Matches 273;	Conservative 126;	Mismatches 362;	Indels 135;	Gaps 211;
Qy	3	GRKDVYELLQNGASVQARDGGLIPLHNACSFHGAEEVNNLLRHGADPNARDNNNYTPL	62		
Db	85	GHEVVSSELLQREANVDAATKGNFTALHSLAGQAEVKKVLTNGANVNAOSQNGFTPL	144		
Qy	63	HEAAIKGKIDYICVILQHGAEPTIERTDGTALDIA----	114		
Db	145	YMAAENHLEVVYFLDNGASOSLATEGFTPLAVALQGHQDVVSLLENDTKGVRLP	204		
Qy	115	LESARSGNEEKMMALLTPLNYNCHASDCRSTPLHAAAGYNRVKIVQLLHGADVHAK	174		
Db	205	ALHIAARKDDTKAALLQNDNNADVESKSGFTPLHIAAHYGNINVTALLNRAAAVDFT	264		
Qy	175	DKGDLVPLHNACSYGHEVTELLVKGACVNMADLWFOFTPLHEAASKNRVEVCSLLLSVG	234		
Db	265	ARNDTITPLHVASKRGNNANMVKLLDORGAIDAKTRDGLTPLHCGARSGHEQVEMLLDRA	324		
Qy	235	ADPTLLNCHNKSADIDLAPTPOLKERLAYEFKGLSLQAAAREADVTRIKKHLSELMVNFKH	294		
Db	325	A--PILSKTKN-----GLSPLHMATQGD-----HLNCVQLLQLH	356		
Qy	295	-----POTHETALHCAASPYPRKKQICEILLRKGANINKEKTEFLPLHVASEKAHN	347		
Db	357	NVPVDDVTNYDITLHVAACHGHYK---YAKVLLDKKANPNKAKALNGFTPLHIACKKNRI	413		
Qy	348	DVEVVVKHEAKYNALDNLGOTSLHRAAYCGHLOTCLLLSYSCDNPNIISLOGFTALQMG	407		
Db	414	KVNEILLKKGASIQAVTESGLTPIHVAAPMGVHNIVYSQLMHGASNTNVRGETALHWA	473		
Qy	408	-----NENYQQLLQEGISLGNSEADRRQ---LLEAAKAGDVETVKKLCTV-OSVNCRDIEGR	459		
Db	474	ARSGQAEVVRYLVODCAQVEAKAKDDQTPLHISARLUGKADIVQOQLLQOQASPNAAATSG-	532		
Qy	460	QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEYVALLVKHCVV	519		
Db	533	-YTPLHLSAREGHEDVAFFLLDHGASLSITTKKGFPTPLHVAAYKGLVEANLLLQKASCP	591		
Qy	520	NVADLWKFTPLHEAAAKGYEICKLILQHGADPTKKNRGDNTPLDL-VKDGDTDIDQLL-	577		
Db	592	DAAGKSGSLTPLHVAHYDNQVKVALLLLDQASPHAAKNGYTPHLIAAKNQMDIATLL	651		
Qy	578	-----RGDAALLDAAKGCIARVKKLSSPD--NVNCRDTQGRHSHTPLHLAAGYNN	625		
Db	652	EYGADANAVTROCISAVHLLAAEGHVDWVSLILGRNANVLSKSG--LPLHLAAQEDR	709		

```

Db 397 RYMEILLKHGASIQAVTESGLTPIHVAFAFMGHVNIQSOLMHHGASPNNTNVRGETALHMA 456
QY 408 -----NENVOQLLOEGISLGNSEADRO--LLEAAKAGDVEVVKLCYV-QSVNCRDIEGR 459
Db 457 ARSGAEVVRVYVQGAOQVEAKAKDDQTPHISARLGKADIVQOQLLOQASPNAAATTS- 515
QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAACSYGHYEVAELLYKHGAVV 519
Db 516 -YTPLHLAAREGHEDVAFAFLDHGASISITTKGFTPLHVAAKYKGLVALLLOKQAS 574
QY 520 NVADLWKEFTPLHEAAKGYEICKLLOHGADPTKKNRDNPTPLDL-VKGDGTDIQLL- 577
Db 575 DAAGKSLGTLPLHVAHYDNQVALLLLDQASPHAAAKNGYTPHLHIAAKNQMDIATSL 634
QY 578 -----RGDAALLDAAKGCCLARVKKLSPD-VNCRDQGRHSTPLHAAAGYNN 625
Db 635 EYGADANAVTQGIASVHLAAQEGHDMVLSLLSRNANVLSNKG--LTPHLHAAQEDR 692
QY 626 LEVAEYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
Db 693 VNVAEVLVNOGAHVDAQTKMGYTPHVGCHYGNIKIVNFFLLOHSAKYNKNGYTAHQ 752
QY 686 AAKGRGTQLCALLAHGADPTLKNQEGTDLIVS-----ADDVSALLTAAMPSPALP 738
Db 753 AAQOGHTHIINVLQNNASPNELTVNGTALAIARRLGYSVVDTLKVVTTEEIMTTIT 812
QY 739 SCYK---POVLNGVRSPGATADALSSPSPPSSLSAASLNLNLS-GSFSELSVVSSSGT 794
Db 813 EKHKNMVPETMNEVLD-----MSDDEVKASAEKLSGGEY-----ISDGE 853
QY 795 EGASSLEKKEVPGVDFSIQFVRLNGLHLMDFEREQITLDVLVEMGHKELKEIG 850
Db 854 EG-----EDAITGDTDKYLGPODLKELG 876

```

## RESULT 5

```

T42715
ankyrin 3, splice form 3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42715
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo, I.
J. Cell Biol. 130, 313-330, 1995
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42715
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1940 <PEP>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 834/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

```

```

Query Match 15.1%; Score 871.5; DB 2; Length 1940;
Best Local Similarity 28.3%; Pred. No. 5e-46;
Matches 320; Conservative 159; Mismatches 464; Indels 189; Gaps 35;

```

```

QY 3 GRKDVVEYLLQNGASVOARDGGLIPLHNAACSYGHAEVYNLLLRHGADPNARDNMNVTPL 62
Db 68 GHEVVEYSELLEAREANVDAATKKGNTALHIAAGQAEVVKLVLTNGANVNAQSQNGFTPL 127
QY 63 HEAAITKGTIDVCIVLLOHCAEPTIRNTDRTALDLA-----DPSAKAVLTGEYK----KDE 114
Db 128 YMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQOQHDOQVVSLLLENDTKGKVRLP 187
QY 115 LLESARSGNEERKWMALLTPLNVNCHASDGRKSTPLHAAAGYNRVKIVQLLQHGADVHAK 174

```

```

Db 188 ALHIAARDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNINVTATLLNRAAAVDFT 247
QY 175 DKGDVPLHNAACSYGHYEVTETLLVKKHACVNAMDWQFPTPLHEAAKRNVEVCSLLSYG 234
Db 248 ARNDITPLHVASKRGNANWVKILLDRGAKIDAKTRDGLTPLHCGARSHEQVVEMLLDRS 307
QY 235 ADPTLLNCHNKAIDLAPTPQLKELAYEFKSHLSLQAAAREADVTRIKKHLSLEMNFKH 294
Db 308 A-PILSKTKN-----GLSPLHMAQGD-----HLNCVQLLLOH 339
QY 295 -----PQTHETALHCAASPYPKRKOICELLRLKGNINEKTEFTPLHVAASEKAHN 347
Db 340 NVPVDDVTNDYTLHVAHAAHCHYK---VAKVLLDKKASPNKALNGFTPLHIACKNRI 396
QY 348 DVVEVVKHAEKVNALDNLGQTSIHRAAYCGHLOTCRLLSYGCDPNIISLQGTALOMG 407
Db 397 RYMEILLKHGASIQAVTESGLTPIHVAFAFMGHVNIQSOLMHHGASPNNTNVRGETALHMA 456
QY 408 -----NENVOQLLOEGISLGNSEADRO--LLEAAKAGDVEVVKLCYV-QSVNCRDIEGR 459
Db 457 ARSGAEVVRVYVQGAOQVEAKAKDDQTPHISARLGKADIVQOQLLOQASPNAAATTS- 515
QY 460 QSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGGLVPLHNAACSYGHYEVAELLYKHGAVV 519
Db 516 -YTPLHLAAREGHEDVAFAFLDHGASISITTKGFTPLHVAAKYKGLVALLLOKQAS 574
QY 520 NVADLWKEFTPLHEAAKGYEICKLLOHGADPTKKNRDNPTPLDL-VKGDGTDIQLL- 577
Db 575 DAAGKSLGTLPLHVAHYDNQVALLLLDQASPHAAAKNGYTPHLHIAAKNQMDIATSL 634
QY 578 -----RGDAALLDAAKGCCLARVKKLSPD-VNCRDQGRHSTPLHAAAGYNN 625
Db 635 EYGADANAVTQGIASVHLAAQEGHDMVLSLLSRNANVLSNKG--LTPHLHAAQEDR 692
QY 626 LEVAEYLLOHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
Db 693 VNVAEVLVNOGAHVDAQTKMGYTPHVGCHYGNIKIVNFFLLOHSAKYNKNGYTAHQ 752
QY 686 AAKGRGTQLCALLAHGADPTLKNQEGTDLIVS-----ADDVSALLTAAMPSPALP 738
Db 753 AAQOGHTHIINVLQNNASPNELTVNGTALAIARRLGYSVVDTLKVVTTEEIMTTIT 812
QY 739 SCYK---POVLNGVRSPG-----ATADALSS-----GP----- 763
Db 813 EKHKNMVPETMNEVLDMSDDGDKCWFKIPKVOEVLVKSDEDAITGDTKYLGPODLKEL 872
QY 764 ---SSPSSLSAASLNLNLSGFSLSVSVS-----SSGTEGASSLEKKEVPGVDFSIQF 815
Db 873 GDSLPAEGYVGFSLGARSASLSRFSRSRSTYTLNRSYARDSMMIEELLVPSKEQHLT-F 931
QY 816 VRNLGLEHLMDFEREQITLDVL-----VEMGHKELKEITGINAYG-----HRHKL 861
Db 932 TREDFDSLRH-YSWAADTLDNVNLVSSPVHSGF--LVSEFMDARGSGMRSGRHGMRII 988
QY 862 -----KGVRLISGOGLNPVLTTLNTSGSTLIDLSPDDKEFQS-VEEEMQSTV 910
Db 989 IPRKCTAPTRITCLVLRHKLANPPMVEGELSLRVLVEMGPAGAQFLGPVTEIP--- 1045
QY 911 REHRDGGHAGFPNRYNLIKQVKNKLEW---RYTHRRKEVSE-----ENHNHANERML 963
Db 1046 -----HFGSRGKRELIIVLRSENGETWKEHQFDSKEDLAELLNGDDELDSPLE- 1097
QY 964 FHGSPFNALIHKGDFERHAYIGGMEGAGIYFAENS---SKSNQYVYGGGG 1012
Db 1098 --GTRICRIITIKDPQ-----YFAVWSRIKQESNQ--IGPEGG 1132

```

## RESULT 6

```

S37431
ankyrin 2, neuronal long splice form - human
N:Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythro
N:Contains: ankyrin 2, short form

```

C:Species: Homo sapiens (man)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999  
 A:Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569  
 R:Chan, W.  
 submitted to the EMBL Data Library, September 1993  
 A:Reference number: S37431  
 A:Accession: S37431  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-3924 <CHA>  
 A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288  
 R:Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.  
 J. Cell Biol. 114, 241-253, 1991  
 A:Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a  
 A:Reference number: A39643; MUID:91302466; PMID:18330053  
 A:Accession: A39643  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2077 <Orl>  
 A:Cross-references: GB:X56957  
 A:Accession: B39643  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1443,3585-3924 <OTT>  
 A:Cross-references: EMBL:X56958  
 R:Tse, W.T.; Menniger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,  
 Genomics 10, 858-866, 1991  
 A:Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.  
 A:Reference number: A40334; MUID:92009921; PMID:1833308  
 A:Accession: A40334  
 A:Molecule type: DNA  
 A:Residues: 463-474, 'PE', 477-495 <TSE>  
 A:Cross-references: GB:M37123; NID:9178647; PIDN:AAA62828.1; PID:9178648  
 R:Chan, W.; Kordeili, E.; Bennett, V.  
 J. Cell Biol. 123, 1463-1473, 1993  
 A:Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and se  
 A:Reference number: A49462; MUID:94075409; PMID:8253844  
 A:Accession: A49462  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-3924 <RES>  
 A:Cross-references: EMBL:Z26634; NID:9406287; PIDN:CAA81387.1; PID:9406288  
 C:Genetics:  
 A:Gene: GDB:ANK2  
 A:Cross-references: GDB:127607; OMIM:106410  
 A:Map position: 4q25-4q27  
 C:Superfamily: ankyrin; ankyrin repeat homology  
 C:Keywords: alternative splicing  
 F:2-3924/Product: ankyrin 2, long form #status predicted <MAT>  
 F:2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>  
 F:63-95/Domain: ankyrin repeat homology <AN01>  
 F:96-128/Domain: ankyrin repeat homology <AN02>  
 F:129-161/Domain: ankyrin repeat homology <AN03>  
 F:162-190/Domain: ankyrin repeat homology <AN04>  
 F:191-223/Domain: ankyrin repeat homology <AN05>  
 F:232-264/Domain: ankyrin repeat homology <AN06>  
 F:265-297/Domain: ankyrin repeat homology <AN07>  
 F:298-330/Domain: ankyrin repeat homology <AN08>  
 F:331-363/Domain: ankyrin repeat homology <AN09>  
 F:364-396/Domain: ankyrin repeat homology <AN10>  
 F:397-429/Domain: ankyrin repeat homology <AN11>  
 F:430-462/Domain: ankyrin repeat homology <AN12>  
 F:463-495/Domain: ankyrin repeat homology <AN13>  
 F:496-528/Domain: ankyrin repeat homology <AN14>  
 F:529-561/Domain: ankyrin repeat homology <AN15>  
 F:562-594/Domain: ankyrin repeat homology <AN16>  
 F:595-627/Domain: ankyrin repeat homology <AN17>  
 F:628-660/Domain: ankyrin repeat homology <AN18>  
 F:661-693/Domain: ankyrin repeat homology <AN19>  
 F:694-726/Domain: ankyrin repeat homology <AN20>  
 F:727-759/Domain: ankyrin repeat homology <AN21>  
 F:760-792/Domain: ankyrin repeat homology <AN22>  
 F:793-825/Domain: ankyrin repeat homology <AN23>

Query Match 14.9%; Score 860.5; DB 2; Length 3924;  
 Best Local Similarity 29.1%; Pred. No. 7.5e-45;  
 Matches 297; Conservative 147; Mismatches 404; Indels 174; Gaps 29;  
 QY 3 GRKDVEYLLONGASVOARDGGLIPLHNACSGHAEVWNLRLRHGADPNARDNNWYTP 62  
 DB 75 GHVGLVQELLGRGSSVDSATKKNTALHIALAGQAEVWVVLKREGANINAQSONGFTPL 134  
 QY 63 HEAAIKGKIDVICVILLOHGAFTIRNTDGT---ALDLADPSAKAVLTGEYKDEL--- 115  
 DB 135 YMAAOENHIDVVKYLLENGANOSTATEDGFTPLAVALQOQHNOAVALLLENDTKGKVRP 194  
 QY 116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLAAGYNRVKIVQL 163  
 DB 195 ALHIAARKDDTKSAALLLQ---NDHNADVQSKMVMNRTTESGFTPLHIAAHYGNVATL 251  
 QY 164 LLQHGADVHAOKGDLVPLHNACSYGHVTELLVKKGACVNMALWQFTPLHEAASKNR 223  
 DB 252 LLNRGAADVFTARNGITPLHVASKRGNTNMVKLLLDGRGQ IDAKTRDGLTPLHCAARSGH 311  
 QY 224 VEVCSLLLSYGADPTLLNCHNKSALDLAPTQLKERLAYEFKHSLLQAAAREADVTRIKK 283  
 DB 312 DOVVELLLERGA-PLLARTKN-----GLSPHMAAQGDHVECVK 349  
 QY 284 HLSLEMVNFKHP-----QTHETALHCAASPYPRKQICEILLRKGANINEKTFEITPL 338  
 DB 350 HL----LOHKAPVDVDTLYLTALHVAA---HCGHYRVTKLLDKRANPNARALNGFTPL 402  
 QY 339 HVASEKAHNDVVVVVKEAKVNALDNLGQTSLHRAAYCGHLOTCRLLSLLSGCDPNISL 398  
 DB 403 HIACKNKRIKVMELLYKYGASIOAITESGLTPIHVAAPMCHLNIVLLLLONGASPDVTNI 462  
 QY 399 OGFTALQMGNNVQQLLOEGISLGNSEADROLLEAAKAGDVETVKKLCITGVQSNCRDIEG 458  
 DB 463 RQETALHM-----AARAGQVEVVR--CLLRNGALVDARA 494  
 QY 459 R-QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVPLHNACSYGHVYVAELLVKHGA 517  
 DB 495 REEQTPHLIASRLGKTEIVQLLQHMHPDAATTNGYTPHLISAREGOVDVASVLEAGA 554  
 QY 518 VYVNVADLWKFPLHFAAAKGYEICKLLLOHCADPTKKNRGDNTPLDLVKGDDTDIDOLL 577  
 DB 555 AHSLATKKGFTPLHVAAYKGLDVAKLLLRRAADSAGKNGLTPLHVAHYDN----- 608  
 QY 578 RGDAAALLDAAKGLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGA 637  
 DB 609 -----QKVALLLLEKASP-----HATAKNGYTPHLIAAKKNQMIASLLNYGA 653  
 QY 638 DYNAQDKGLIPLHNAASYGHVDVVAALLIKYNACVNATDKNAFTPLHEAAQGRTOCAL 697  
 DB 654 ETNIVTKQGVTPHLIASQEGHTDMVTLLLDKGANIHMSTKSGTSLHLAAQEDKYNVADI 713  
 QY 698 LLAHGADPTLKNOEGOTPLDIV-----SADDVSALLTAAMPSPA-LPSCVKP----- 743  
 DB 714 LTKHGADQDAHTKGLTTPLIIVACHYGNVKNWNLKQGANVNAKTKNGTTPHQAQOQH 773  
 QY 744 -OVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSFSELSSVSSSGTEGASSLEK 802  
 DB 774 THIIINVLLQHGAKPNATTA--NGNTALAIKRLGYI--SVVDTLKVVTEEVTTTTTITE 829  
 QY 803 KEVPGVDFSIQFVRNMLGHEHMDIFERE---QITLDVLVEMGHKELKEIGNAYCHRRHK 859  
 DB 830 KHKLNVPETMTE-----VLDVSDERGGDDTMTGDGGEYLRPEDLDELGSDDS----- 874  
 QY 860 LIKGVRLISGO--OGLNPLYLTNTSGSTILIDILSPDKFQSVSEEMQSVREHROGG 917  
 DB 875 -----LPSSQFLDGMN-ILYRLSLEGG-----RSDLSRFSSSDRSHTLSHASYLRD-- 918  
 QY 918 HAGGIENRYNLIKQKVCN-KKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAITHK 976  
 DB 919 --SAVMDSDSVIPSHQVSTLAKFAERNRYLSWGNTENLDNVA-----LSSSP-----IHS 966

```

Qy 977 GF 978
Db 967 GF 968

RESULT 7
B35049
N:Alternate names: ankyrin 1, erythrocyte splice form 3 - human
N:Contains: ankyrin 2.1, erythrocyte; ankyrin-R
C:Species: Homo sapiens (man)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C:Accession: B35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370; PMID:1689849
A:Accession: B35049
A:Status: preliminary
A:Molecule type: mRNA
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-1513,1676-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 14.6%; Score 844; DB 2; Length 1856;
Best Local Similarity 30.1%; Pred. No. 2,5e-44;
Matches 242; Conservative 121; Mismatches 263; Indels 177; Gaps 21;

Qy 3 GRKDVVEYLLQNGASVOARDGGLPLHNACSGHAEVYNLLRHGADPNARDNNVYPL 62
Db 89 GDEVVRELNVNGANVNAQSQKGTPLYMAAQENHLEVVKFLLCANGQNVATEDGFTPL 148
Qy 63 HEAAIKGKIDVCIVLQHGAEPI-----RNTDGRITDLADPSAKAVUTGEYKDK 113
Db 149 AVAQGHENVVAHLNINYGKGVRLPALHIAARNDDTRTA-----AVLLQNDPNP 199
Qy 114 ELLESARENEEKMMALLTPLLNVNCHASD-----GRKS-----TPHLAAGY 155
Db 200 DVL-SKTG-----FTPLHAAHYENLVNAQALLNRGSSVNFPPQNGITPLHIASR 249
Qy 156 NRKVIQVOLLQHGADVHAKDGLVPLHNACSGHYEVTELLVKKHACVNAQMDLWQFTPL 215
Db 250 GNVIMVRLLDGRGAQITETKTDELTPHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309

```

```

Qy 216 HEASKNEVEVCSLLLSYGADPTLLNCHNKSAIDLAPTLPOLKERLAYEFKGHSLLOAARE 275
Db 310 HMAAGDHLDVCVRLLOYDAE-----ID----- 332
Qy 276 ADVTRIKKHLISLEMVNFKHPQTHETALHCAASPPYKPKQICEILLRKGANINEKTEKFL 335
Db 333 -DIT-----LDHLLPLHVA-----HGHIRVAKVLLDKGAPNSRALNGF 372
Qy 336 TPLHVASEKAHNDVVVVKHEAKVNALDNLGOTSLHRAAYCGHILQTCRLLLSYGCDPNI 395
Db 373 TPLHACKKNHVRVMEILLKGTASIDAVTESGLTPLHVASEMGHLPVKNLLQRCASPNV 432
Qy 396 ISLOGFTALQMG-----NENVOQLLOE-----GISLGNSEADROLLE- 432
Db 433 SNKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDOTPLHCAARIGHTNMVKLLLEN 492
Qy 433 -----AAKAGDVETVKKLCTVO-SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db 493 NANPNLATTAGTTPHIAARESHVETVLALLEKASQACMTKKG--FTPLHVAAYKGVKVR 550
Qy 475 VVEYLLQHGADVHAKDGGGLVPLHNACSGHYEVAELLVKHGAVNVNADLWK-FTPLHEA 533
Db 551 VAELLERDAHPNAAKNGLTPLHVAVHNNLDIVKLLPRGSGPH-SPANNGVYTPLHIA 609
Qy 534 AAKGYEICKLLQHGADPTKKNRDGNTPDL-VKDGDTDIQDLLRGLDAAALLDAKKGCL 592
Db 610 AKQOQVEVARSLLOYGGSANRESVOGVTPLHIAAQEGHAEVALL----- 654
Qy 593 ARVKLLSPDNVNCRDQTGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAODKGLIPLHN 652
Db 655 -----LSKQANGNLGKSG--LTPHLVVAQBGHPVADVLKKGVMVDATTRMGYTPH 707
Qy 653 AASGYHVDVAALLIKYNACVNAATDKWAFPLHEAAQGRQTCALLLAHAGADPTLKNOEG 712
Db 708 ASHYGNIKLVKFLQHQADVNAATKLGYSPLHQAQOQGTDTIVTLLKNGASPNVSSDG 767
Qy 713 QTPLDL-----VSADVDVSALLT 729
Db 768 TTPLAIKRLGYISVTDVLKVV 790

RESULT 8
A35049
N:Alternate names: ankyrin 1, erythrocyte splice form 2 - human
N:Contains: ankyrin 2.1, erythrocyte; ankyrin-R
C:Species: Homo sapiens (man)
C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370; PMID:1689849
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAM>
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MA2>
F:2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>

```

F:271-303/Domain: ankyrin repeat homology <AN08>  
F:304-336/Domain: ankyrin repeat homology <AN09>  
F:337-369/Domain: ankyrin repeat homology <AN10>  
F:370-402/Domain: ankyrin repeat homology <AN11>  
F:403-435/Domain: ankyrin repeat homology <AN12>  
F:436-468/Domain: ankyrin repeat homology <AN13>  
F:469-501/Domain: ankyrin repeat homology <AN14>  
F:502-534/Domain: ankyrin repeat homology <AN15>  
F:535-567/Domain: ankyrin repeat homology <AN16>  
F:568-600/Domain: ankyrin repeat homology <AN17>  
F:601-633/Domain: ankyrin repeat homology <AN18>  
F:634-666/Domain: ankyrin repeat homology <AN19>  
F:667-699/Domain: ankyrin repeat homology <AN20>  
F:700-732/Domain: ankyrin repeat homology <AN21>  
F:733-765/Domain: ankyrin repeat homology <AN22>  
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 14.6%; Score 844; DB 2; Length 1880;  
Best Local Similarity 30.1%; Pred. No. 2.5e-44;  
Matches 242; Conservative 121; Mismatches 263; Indels 177; Gaps 21;

QY 3 GRKDVVEYLQNGASVQARDGGGLIPLHNACSFHGAEEVNVNLLRHGADPNARDNNWYTP 62  
DB 89 GQDEVVRELNVNGANVNAQSGKFTPLMAAENHLEVVKFLLENGANQNVATEDGETPL 148  
QY 63 HEAAIKGKIDVICVILQHGAEPTI-----RNTDORTALDLADPSAKAVLTGEYKKD 113  
DB 149 AVALQOQHENVVAHLINTGKGVRLPALHIAARNDDRTA-----AVLLQNDPNP 199  
QY 114 ELLESARSNEEKMMALLTPLNVNCHASD-----GRKS-----TPLHIAAGY 155  
DB 200 DVL--SKTG-----FTPLHIAAHYENLVNAQLLNNGSSVNPFGITPLHIAARR 249  
QY 156 NRKIVVQLLQHGADVHAKDGDVPLHNACSYGHYEVTELLVKGACVNAWMDLMOFTPL 215  
DB 250 GNVIVVRLLDGQAETKTDELTPHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309  
QY 216 HEAAKNRVEVCSLLSYGADPTLLNCHNKSATDLAPTQLKERLAYEFKGHSLLQAARE 275  
DB 310 HMAAQGDHLDCVRLLOYDE-----ID----- 332  
QY 276 ADVTRIKHLSLEMVNFKHPOTHETALHCAAAAPYKPKQICELLRKGANINEXTKKFL 335  
DB 333 -DIT-----LDHLTPLVHAA---HCGHHRVAKVLLDKGAKPNRSLANGF 372  
QY 336 TPLHVAASEKANDVVEVVKHAKYNALDNLGQTSILHRAAYCGHLQTCRLLLSYCDPNI 395  
DB 373 TPLHIAACKNHVRVMBELLTKTASIDAVTESGLTPLHVASFHGLPIPVKNLLQORGASPNV 432  
QY 396 ISLQGTFTALQMG-----NENVQOLIQE-----GTSLGNSEADRLQLE- 432  
DB 433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPPLHCAARIGHTNMVKLLLEN 492  
QY 433 -----AAKAGDVETVKKLCTVQ- SVNCRDIEGROSTPLHPFAAGNVRVS 474  
DB 493 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAYGKYR 550  
QY 475 VVEYLLQHGADVAKDKGLVPLHNACSYGHYEVTELLVKGACVNAWMDLMOFTPLH 533  
DB 551 VAELLERDAPHPNAAGKNGLTPLHVAVHNNLDIVKLLPLRGSGSPH-SPAWNCGYTPH 609  
QY 534 AAKGKEITCKLLQHGADPTPKNRDGNTPDL- VKDGDGTDIDQLLRGDAALLDAKKGCL 592  
DB 610 AKQNVETVARSLLQYGGSSANESVQGVTPPLHIAAQEGHAEWALL----- 654  
QY 593 ARVKKSSPDNVNCRDQGRHSHTPLHIAAGYNVLEVAEYLLQHGADVNAQDKGGLTPLHN 652  
DB 655 -----LSKQANGNLNKGSG--LTPLHLVAQEGHPVADVILKHHGVMDVATTRMGYTP 707  
QY 653 AASGYHVDVAALLIKYNACVNAWMDKWAFTPLHEAAQKGRTOCALLLAHGADPTLKNQEG 712  
DB 708 ASHYGNLKVLFLLQHQADVNAKTKLGYSPHQAQOOGHTDIVTLLLLKNGASPNVSSDG 767

QY 713 QTPDLI-----VSADDDVSALLT 729  
DB 768 TTPLAIAKRLGILSVTDVLKQVT 790

RESULT 9  
SJHUK

N:Ankyrin 1, erythrocyte splice form 1 - human  
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R  
N:Contains: ankyrin 2.2  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 22-Jun-1999  
C:Accession: S08275; A33219; PC2220; A35443  
R:Lux, S.E.; John, K.M.; Bennett, V.  
Nature 344, 36-42, 1990  
A:Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure  
A:Reference number: S08275; MUID:90158830; PMID:2137557  
A:Accession: S08275  
A:Molecule type: mRNA  
A:Residues: 1-1881 <LUI>  
A:Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702  
A:Accession: A33219  
A:Molecule type: protein  
A:Residues: 2-7, 'X', '9-17', 'X', '19-20', 'T', '22-30', '733-749', 'A', '751-753', '828-833', 'X', '835-855', 'X', '1367', '1383-1427', '1601-1630', '1686-1698', 'P', '1700', '1763-1772' <LUX>  
A:Note: 845-Arg and 1392-Thr were also found  
R:Hermann, J.; Barel, M.; Frade, R.  
Biochem. Biophys. Res. Commun. 204, 453-460, 1994  
A:Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membr  
A:Reference number: PC2220; MUID:95071348; PMID:7526850  
A:Accession: PC2220  
A:Molecule type: protein  
R:Davis, L.H.; Bennett, V.  
J. Biol. Chem. 265, 10589-10596, 1990  
A:Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchang  
A:Reference number: A35443; MUID:90285190; PMID:2141335  
A:Accession: A35443  
A:Molecule type: protein  
A:Residues: 'X', '5', 'X', '7-12', '403-417', 'X', '419-422', 'H', '424', 'LQ', '797-800', 'L', '802-814', '862-8  
C:Genetics:  
A:Gene: GDB:ANK1; ANK  
A:Cross-references: GDB:118737; OMIM:182900  
A:Map position: 8p11.2-8p11.2  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing; phosphoprotein  
F:2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>  
F:2-1512,1675-1881/Product: ankyrin 2, erythrocyte #status predicted <MAT2>  
F:2-827/Domain: 89K #status predicted <DOM1>  
F:2-827/Region: anion exchange protein binding  
F:44-76/Domain: ankyrin repeat homology <AN01>  
F:77-109/Domain: ankyrin repeat homology <AN02>  
F:110-142/Domain: ankyrin repeat homology <AN03>  
F:143-171/Domain: ankyrin repeat homology <AN04>  
F:172-204/Domain: ankyrin repeat homology <AN05>  
F:205-237/Domain: ankyrin repeat homology <AN06>  
F:238-270/Domain: ankyrin repeat homology <AN07>  
F:271-303/Domain: ankyrin repeat homology <AN08>  
F:304-336/Domain: ankyrin repeat homology <AN09>  
F:337-369/Domain: ankyrin repeat homology <AN10>  
F:370-402/Domain: ankyrin repeat homology <AN11>  
F:403-435/Domain: ankyrin repeat homology <AN12>  
F:436-468/Domain: ankyrin repeat homology <AN13>  
F:469-501/Domain: ankyrin repeat homology <AN14>  
F:502-534/Domain: ankyrin repeat homology <AN15>  
F:535-567/Domain: ankyrin repeat homology <AN16>  
F:568-600/Domain: ankyrin repeat homology <AN17>  
F:601-633/Domain: ankyrin repeat homology <AN18>  
F:634-666/Domain: ankyrin repeat homology <AN19>  
F:667-699/Domain: ankyrin repeat homology <AN20>  
F:700-732/Domain: ankyrin repeat homology <AN21>  
F:733-765/Domain: ankyrin repeat homology <AN22>  
F:766-798/Domain: ankyrin repeat homology <AN23>

F:828-1382/Domain: 62k #status predicted <DOM2>  
F:828-1382/Region: spectrin binding  
F:1383-1881/Domain: 55k #status predicted <DOM3>

Query Match 14.6%; Score 842; DB 1; Length 1881;  
Best Local Similarity 29.9%; Pred. No. 3.4e-44;  
Matches 240; Conservative 122; Mismatches 264; Indels 177; Gaps 20;

```
QY 3 GRKDVVEYLQNGASVOARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 62
Db 89 GODEVRELNVYGANVNAOSQKGTPLYMAAQENHLEVVKKLENGANQNVATEDGFTPL 148
QY 63 HEAAIKGKIDVICVILLOHGAETPI-----RNTDGRALTDLADPSAKAVLTGEYK 113
Db 149 AVALQOQHENVVHLYNGTKGVRLPALHIAARNDDTRTA-----AVLLQNDPNP 199
QY 114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLHLAGY 155
Db 200 DVL--SKTG-----FTPLHIAHYENLVNVAQLLNRGASVNFPTQNGITPLHIA 249
QY 156 NRKIVOLLQHGADVHAKDGDLVPLHNACSYGHVEVTELLVKGACVNAIDLWQFTPL 215
Db 250 GNVIMVRLLDRCAGIETRTKDELTPHCAARNGHVRISIELLDHGAPIQAKTKNGLSPI 309
QY 216 HEAAKRNVEVCSLLSYGADPTLLNCHNKSADIDLAPTQPKERLAYEFKSHSLQAARE 275
Db 310 HMAAQGDHLDCVRLLOQYNAE-----ID----- 332
QY 276 ADVTRIKKHLNLEWVNFKHPQTHETALHCAASPYKPKKQICELLRLKGANINEKTEFL 335
Db 333 -DIT-----LDHLTPLHVA-----HCGHHRVAKVLLDKGAKPNRSLNGF 372
QY 336 TPLHVAEKAHNDVVEVVKHKAQVNLNGLQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
Db 373 TPLHIAKKNHVRVMEILLTKGTASIDAVTESGLTPLHVASFMGHLPVKNLQORGASPNV 432
QY 396 ISLQGFALOMG-----NENVQQLQE-----GISLGNSEADRLLE- 432
Db 433 SNVKVETPLHMAARAGHTEVAKVLLQNKAKVNAKAKDDQTPHCAARIGHTNMVKLLLEN 492
QY 433 -----AAKAGDVETVKLCITVQ--SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db 493 NANPNLATTAGHTPLHIAAREGHVEITVALLLEKESQACMTKG--FTPLHVAAKYGV 550
QY 475 VVEYLLOHGADVIAKDKGLVPLHNACSYGHEVVAELLVKHGAVNVADLWK-FTPLHEA 533
Db 551 VAELLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLPRGGSPH-SPAWNGYTPHIA 609
QY 534 AAKGKYEICKLLQHGADPTKKNRDNTPDL-LVKDGTDIQDLRGDAALLDAARKGCL 592
Db 610 AKQNOVEVARSLLOYGGSANESVQGVTPPLHIAAQEGHAEWVALL----- 654
QY 593 ARVKKLSSPDNNCRDQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAQDKGLIPLHN 652
Db 655 -----LSKQANGNLNKG--LTPHLVAGEGHVPVADVLLKHGVNVDAITRNGYPLHV 707
QY 653 AASYGHVDVAALLIKYNACVNATDKWAFVTPHFAAGKRTQICALLHAGADPTLNQEG 712
Db 708 ASHYGNKIVKFLLOHQAADVNAKTILGYSPLHQAQOQHTDIVTLLKNGASPNVSSDG 767
QY 713 QTPLDL-----VSADDVSALLT 729
Db 768 TPLAIAKRILGISVTDVLLKVT 790
```

RESULT 10

S37771

ankyrin, erythrocyte - mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-May-1994 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999

C:Accession: S37771

R:Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.  
J. Biol. Chem. 268, 9533-9540, 1993

A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found  
A:Reference number: S37771; MUID:93252825; PMID:8486643

A:Accession: S37771

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1848 &lt;BIR&gt;

A:Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

F:48-80/Domain: ankyrin repeat homology &lt;AN01&gt;

F:81-113/Domain: ankyrin repeat homology &lt;AN02&gt;

F:114-146/Domain: ankyrin repeat homology &lt;AN03&gt;

F:147-175/Domain: ankyrin repeat homology &lt;AN04&gt;

F:176-208/Domain: ankyrin repeat homology &lt;AN05&gt;

F:209-241/Domain: ankyrin repeat homology &lt;AN06&gt;

F:242-274/Domain: ankyrin repeat homology &lt;AN07&gt;

F:275-307/Domain: ankyrin repeat homology &lt;AN08&gt;

F:308-340/Domain: ankyrin repeat homology &lt;AN09&gt;

F:341-373/Domain: ankyrin repeat homology &lt;AN10&gt;

F:374-406/Domain: ankyrin repeat homology &lt;AN11&gt;

F:407-439/Domain: ankyrin repeat homology &lt;AN12&gt;

F:440-472/Domain: ankyrin repeat homology &lt;AN13&gt;

F:473-505/Domain: ankyrin repeat homology &lt;AN14&gt;

F:506-538/Domain: ankyrin repeat homology &lt;AN15&gt;

F:539-571/Domain: ankyrin repeat homology &lt;AN16&gt;

F:572-604/Domain: ankyrin repeat homology &lt;AN17&gt;

F:605-637/Domain: ankyrin repeat homology &lt;AN18&gt;

F:638-670/Domain: ankyrin repeat homology &lt;AN19&gt;

F:671-703/Domain: ankyrin repeat homology &lt;AN20&gt;

F:704-736/Domain: ankyrin repeat homology &lt;AN21&gt;

F:737-769/Domain: ankyrin repeat homology &lt;AN22&gt;

F:770-802/Domain: ankyrin repeat homology &lt;AN23&gt;

Query Match

14.6%; Score 839.5; DB 2; Length 1848;

Best Local Similarity 29.5%; Pred. No. 4.7e-44;

Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;

```
QY 3 GRKDVVEYLQNGASVOARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 62
Db 93 GODEVRELNVYGANVNAOSQKGTPLYMAAQENHLEVVKKLENGANQNVATEDGFTPL 152
```

```
QY 63 HEAAIKGKIDVICVILLOHGAETPI-----RNTDGRALTDLADPSAKAVLTGEYK 113
Db 153 AVALQOQHENVVHLYNGTKGVRLPALHIAARNDDTRTA-----AVLLQNDPNP 203
```

```
QY 114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLHLAGY 155
Db 204 DVL--SKTG-----FTPLHIAHYENLVNVAQLLNRGASVNFPTQNGITPLHIA 253
```

```
QY 156 NRKIVOLLQHGADVHAKDGDLVPLHNACSYGHVEVTELLVKGACVNAIDLWQFTPL 215
Db 254 GNVIMVRLLDRCAGIETRTKDELTPHCAARNGHVRISIELLDHGAPIQAKTKNGLSPI 313
```

```
QY 216 HEAAKRNVEVCSLLSYGADPTLLNCHNKSADIDLAPTQPKERLAYEFKSHSLQAARE 275
Db 314 HMAAQGDHLDCVRLLOQYNAE-----ID----- 336
```

```
QY 276 ADVTRIKKHLNLEWVNFKHPQTHETALHCAASPYKPKKQICELLRLKGANINEKTEFL 335
Db 337 -DIT-----LDHLTPLHVA-----HCGHHRVAKVLLDKGAKPNRSLNGF 376
```

```
QY 336 TPLHVAEKAHNDVVEVVKHKAQVNLNGLQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
Db 377 TPLHIAKKNHVRVMEILLTKGTASIDAVTESGLTPLHVASFMGHLPVKNLQORGASPNV 436
```

```
QY 396 ISLQGFALOMG-----NENVQQLQEISLGNSEAD-----ROLLE- 432
Db 437 SNVKVETPLHMAARAGHTEVAKVLLQNKAKVNAKAKDDQTPHCAARIGHTNMVKLLLEN 496
```

```
QY 433 -----AAKAGDVETVKLCITVQ--SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db 497 GASPNIATTAGHTPLHIAAREGHVEITVALLLEKESQACMTKG--FTPLHVAAKYGV 554
```







Best Local Similarity 26.1%; Pred. No. 2.1e-42;		Matches 301; Conservative 156; Mismatches 418; Indels 279; Gaps 36;	
Qy	3	GRKDVVEYLLQNGASVQARDGGGLPLHNAACSGHAEVWVLLLRHGADPNARDNNVTPL 62	
Db	77	GHSEVVRRLIRQAQVDAATKRGNTALHIAAGSLIVILVENGANVNVOSVNGFTPL 136	
Qy	63	HEAATKGRKIDVCI-VLLOHGAEPTRINTDRTALDLPASAKAVITGEYKDELLESARS 122	
Db	137	YMAAQENHEEVVYLLKHGAQALSTEDGFTPLA-----LQQG 176	
Qy	123	NEEKMAALLPLNVNCHASDGRKSTP-LHLAAGYNRVKIVOLLQHGADVHAKDKGLVP 181	
Db	177	HDVVAVILLE-----NDSKGKVRPLPALHIAAKDDTTAATLLQNEHNPVDTSKSFT 230	
Qy	182	LHNACSYGHEYTELLVVRHGCNVAMDLMQFTPLHEAASKRVEVCSLLLSYGADPTLLN 241	
Db	231	LHIAHYGHENVGOLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLSRGA---IID 287	
Qy	242	CHNKSAIDLAPTQLKERLAYEFKSHLSLQAAREADVTRIKKHSLSLMVNVKHPQTHETA 301	
Db	288	SRTKOLL-----TP 296	
Qy	302	LHCAASPYPKRKOICELLRLKGANINEKTEFPLHVASEKAHNDVVEVVKHEAKVN 361	
Db	297	LHCAARSGH---DQVVDLLVQAGPISAKTKNGLAPLHMAAGDHVDAARTLLYHRAVD 353	
Qy	362	--ALDNLGOTS LHRAAYCGHLOTCLRLLSYGCDDPNIISLOGFTALOMGNEN----- 410	
Db	354	DVTVDYL--TPLHVAACHGVHVKALLDSDRSPNSRALNGFTPLHIAACKNRKRWELL 411	
Qy	411	-----VOQLLEGIS-----LQNSEADROLLEAKAGDVETVKKLCITVQSVNCRDIE 457	
Db	412	LKYRAAIEATTESGLTPLHVAAFMGAINIVYLQOAGNPDET----- 456	
Qy	458	GROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNAACSYGHEVAELLVKHGA 517	
Db	457	-RGETPLHAAARANTQDVVRLIRNGAKYDAQARELOTPHLHSLRGNITDIIVILLQAGA 515	
Qy	518	VNVADLWFTPLPEHAAAKGYEICKLLOHGADPTKKNRDNTPDLV-KGDDTDIQL 576	
Db	516	NSNATTRDNYSPHLIAAECQEEVAGILLDHNAKDTLLPKKGTPLHLASKYGNLEVVRL 575	
Qy	577	L--RGDAALLDAACKGLARVKLSSPDVNVNCRDTQGRHS--TPLHAAAGYNNLEVAEYLL 633	
Db	576	LLERGTVP-----DIEGKNQVTPPLHVAHYNNDKVAMLL 610	
Qy	634	OHGADVNAQDKGGLPLHNAASGYHVDVAALLIKYNACVNAITDKWFTPLHEAAQKGRTO 693	
Db	611	ENGASAKAAKNGYTPHLHIAAKKNQMEIASTLLQPKADPNAKSRAGFTPLHLSAQEGHKE 670	
Qy	694	LCALLAHGADPTLKNOEGTPTDLVSAD-----VSALLTAAMPSPALPS 739	
Db	671	ISGLLIENGSDVGAKANGLTAMHLCAQEDHPVPAQILYNNGAIEINSKTNAGVTPPLHV-A 729	
Qy	740	CYKPOVLNGVR---SPGATADALSSGSPSSLSAASSLDNLNLSGFSSELSSVSSSGTEG 796	
Db	730	CHFGQ-LNVKFLVENGADVGEKTRASYTPHQAQOQHNNCVRYLLE-----NC 778	
Qy	797	ASSLEKKEVPGVDFSTQFVNLG-----LEHLMDFEREQITLDLVEMGHKELKEIGIN 852	
Db	779	ASPNEQTATGTPLSIAQ---RLGYVSVVETLRTVTETVITETTVDERYKQPQPEAMN 835	
Qy	853	-----AYGRHKLI-----KGERLISGOOGLNPVLTNTSG 884	
Db	836	ETMFSEDEGAQAAEVAHAHEKDFSDNLTOGLQDSTQGLVMIHTGQLQRQSELENGG 895	
Qy	885	-----STILIDLSPDKFQFQSEEMQS-----TVREHRDG--GHAGGIFNRYNLIKQ 932	
Db	896	AIPKINS-----GMSP-EKEPAKTAPVATSPATNSQSFGIAPRAGSISGQFOQPLH 950	
Qy	933	KVCNKLWERYTHRRKE--VSEENHNHANERMPLPHGS-----PFVNAIHHGF-----DE 980	
<hr/>			
Db	951	GAGPNEDLELVRAQHPIINAGNYDNGGVAMLENHADNVPIGHHTVOTPSFLISFLVDA 1010	
Qy	981	RHAYICGMFGAGIYFAENSSKSNQYVYIGGTCGCPVHKDRSCYICHRQLLCFRTVLGKS 1040	
Db	1011	RGAMRGCRHSRGVRIIVPRKASQ-----PIR-----VTCRY-LRKD 1046	
Qy	1041	FLOFSAMKMAHSP 1054	
Db	1047	-----KLAHPP 1053	
<hr/>			
RESULT 14			
T42691			
hypothetical protein DKFZp434D2328.1 - human (fragment)			
C:Species: Homo sapiens (man)			
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000			
C:Accession: T42691			
R:Bloeker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.			
submitted to the Protein Sequence Database, November 1999			
A:Reference number: Z22230			
A:Accession: T42691			
A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-791 <AAA>			
A:Cross-references: EMBL:AL133087			
A:Experimental source: adult testis; clone DKFZp434D2328			
C:Genetics:			
A:Note: DKFZp434D2328.1			
<hr/>			
Query Match 10.9%; Score 628; DB 2: Length 791;			
Best Local Similarity 27.7%; Pred. No. 2.4e-31;			
Matches 232; Conservative 119; Mismatches 117; Indels 174; Gaps 28;			
Qy	3	GRKDVVEYLLQNGASVQARDGGGLPLHNAACSGHAEVWVLLLRHGADPNARDNNVTPL 62	
Db	14	GOINVVKLLNLGVEIDINVTALHIAAGSLIVILVENGANVNVOSVNGFTPL 73	
Qy	63	HEAATKGRKIDVCI-VLLOHGAEPTRINTDRTALDLPASAKAVITGEYKDELLE----- 116	
Db	74	HFAAATHGALCLLELVNNGADVNIQSKDGKSPHLM-----TAVHGRFTRSTLIQNG 127	
Qy	117	-----ESARSGNEEKMMALLT-----PLN---VNCHA----- 140	
Db	128	ETDCVDDKGNTPHVAARYGHELLINTLITSGADTKACGIHSMFPLHIAALNAHSDCCR 187	
Qy	141	--SDGRK-----STP-----LHLAAGYNRVKIVOLLQHGADV 172	
Db	188	LLSSGKYISIVLSFNEHVLNAGPEIDTPDKFGRVCLHAAAGGNVECIKLQSSGADFH 247	
Qy	173	AKDKGDLVPLHNAACSYGHEYTELLVVRHGCNVAMDLMQFTPLHEAASKRVEVCSLLLS 232	
Db	248	KKDKCGRTPLHYAAANCHFHCTETLVTTGANVNETDDMGRTALHYAAASD----- 297	
Qy	233	YCADPTLL-NCHKSAIDLAPTQLKER---LAYEFKSHLSLQAAREADVTRIKKHSLS 288	
Db	298	MDRNTLILGNAHDNSE-ELERARELKEATLCLEF---LQ----- 335	
Qy	289	MYNKHPOTHE-----TALHCAASPYPKRKOICELLRLKGANINEKTEFUT--PLHVAS 342	
Db	336	--NDANPSIRDEKGYNSIHYAAAYGH---ROCLELLELERTNSGFESDGSATKSPHLAA 390	
Qy	343	EKAHNDVVEVVKHAKVNALENLQGTSLHRAAYCGHLOTCLRLLSYGCDDPNIISLOGFT 402	
Db	391	YNGHQAQLEVLQSLVDLDIRDEKGRGTALDLAFAFKHTECEVALIN-----QGAS 440	
Qy	403	ALOMGNENVQOLLOEGISGNSPADROLLEAAKAGDVETVKKLCITVQSVNCRDIEGROST 462	
Db	441	IFVKNVTNRTPLHASVINGHTLCLLEIAD-----NPEADVVDKAGQ--T 487	
Qy	463	PLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNAACSYGHEVAELLVKHGAHVNA 522	
Db	488	PLMLAVAYGHIDAVSLLEKEANVDTVILGCTALHRTGIMTGHCECVQMLLQEVSLICK 547	

QY 523 DLWFTPLHAAAKGKYETICKLLQHG---ADPTKKNRDGNTPLD--LVKQDGTQIDLL 577  
 Db 548 DSRGRTPPLHYAARGHATWLSLQWALSEEDCCFDKNOGYTPLHWACYNGNCELEVLL 607  
 QY 578 R-----GD-----AALLDAAKKGLARVKKLSLSDNCRDTPQGRHSTPLHLAAG 622  
 Db 608 EOKCRKFTGNFTPLHCAIINDHG--NCASLLGGAIDSSIVSCRDDKGR--TPLHAAAF 663  
 QY 623 YNNLEVAEYLLQHGADVNAQDGGGLIPLHNAASYGHVDVAALLI-KYNACVNATDKWAF 681  
 Db 664 ADHVECLQLLLRHSPVNAVDNSGKTALMMAENQAGAVDILVNSAQADLIVKDKLNT 723  
 QY 682 PLHEAAQKRTQCLCALLAHGADPTL---KNOEGTPTLDVLSADVDVSAALLTAAMPSSA 736  
 Db 724 PUHLACSGHEKCALLILDKIQDESLINEKKNALQTPHLHVAARNGLVVVEELLAKGA 781

RESULT 15  
 S30355  
 alpha-latroinsectotoxin precursor - black widow spider (fragment)  
 C:Species: Latrodectus mactans tredecimguttatus (black widow spider)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Sep-1999  
 C:Accession: S30355; S29864  
 R:Kiyatkin, N.; Dulubova, I.; Grishin, E.  
 A:Title: Cloning and structural analysis of alpha-latroinsectotoxin cDNA. Abundance of a  
 A:Reference number: S30355; MUID:93238678; PMID:8477689  
 A:Accession: S30355  
 A:Molecule type: mRNA  
 A:Residues: 1-1411 <K1Y>  
 A:Cross-references: EMBL:Z14086; NID:g9536; PIDN:CAA78464.1; PID:g9537  
 A:Note: it is uncertain whether Met-26 is the initiator or whether translation is initia  
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology  
 F:Keywords: presynaptic neurotoxin  
 F:533-565/Domain: ankyrin repeat homology <AN1>  
 F:1001-1032/Domain: ankyrin repeat homology <AN2>  
 F:1146-1178/Domain: ankyrin repeat homology <AN3>

Query Match 9.58; Score 545.5; DB 2; Length 1411;  
 Best Local Similarity 23.7%; Pred. No. 9.2e-26;  
 Matches 233; Conservative 136; Mismatches 378; Indels 257; Gaps 33;

QY 11 LLONGASVQARDGGLIPLHNACSFHAEVNVNLLR-----HGADPNARDNNWYTPLHE 64  
 Db 486 LIQGANVSETFELGRGAIHAAASAGWYDCELLNKNLLEKADKN-----GYTPLHI 540  
 QY 65 AAKGKIDVCIIVLQHGAEPTIRNTRDGTALDADPSAKAVLTGEEKKDELLESARGNE 124  
 Db 541 AADSNNKDFVNFLLGNADVNV-----RTKSDLFTP-----LHLAARRDLT 581  
 QY 125 EKMMALLTPLNVNCHASDRKSTPLHAAAGYNRVKIVQLLQHGADVHAKDKGDLVPLHN 184  
 Db 582 DVTQTLDITEIDLNADQKSGFTPLHLSTSTSETAAILIRNTNAVINKSVGLTPLHL 641  
 QY 185 ACSYGHYEVELLVKHCACYNAMDLMQFTPLHEAASKNRVEVCSSLISYCADPTLLNCHN 244  
 Db 642 ATLQNNLSVKLLAGKAGYLNDGANGMTPLHYAAMTGNLEWVDFLL-----N 689  
 QY 245 KSAIDLAPTPOLKE-----RLAYEFKHSLLQ----- 271  
 Db 690 QOYININAAATKEKKWTPLHLAILFKNDVAERLLSDENLNIRLETNGGINPLHASATGN 749  
 QY 272 -----AAREADVTRI-KKHLSELMVNFKHQTHETALHCAASPPKPKQICELLRRK 323  
 Db 750 KQVIELLAKNADVTRLTSGFS-----ALHLGIIG---KNEEIPFLVEK 792  
 QY 324 GANINEKTEFTPLHVASEKANDVVEVVKHEAKVNALDNLGQTSLHRAAYCGHLOTC 383  
 Db 793 GANVNDKNTSGVTPPLHFAAGLKANTFRLLSRGADIKAEIDINSQMPIHEAVSNGHLEIV 852  
 QY 384 RLLSYGCDPNIISLQGTALQWGNENVOQLQE-CISLGNSEADQLLEAAKAGDVETV 442  
 Db 853 RLIEK--DPSLM-----NVKNIRNEYPFYLAVERKRYKDIFDYFVSKD----- 893

QY 443 KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKD----- 490  
 Db 894 -----ANVNEVDHNGNTLLHLFSTGELEVVQFLMQGANFRLKNRERKTFEFLAIE 945  
 QY 491 -----KGLVPLHNACSYHIEVAELLVKGAVNVVADLWKFT 528  
 Db 946 NGRNLIVAFVBEKNKVNLOAHRGKTYLHACDSAKYDKIEIVKYFIEKLNESCE---N 1002  
 QY 529 PLHEAAAKGYEICKLLQ-HGADPTKKNRDGN--TPDLVVKDG-----DTDIDQLLRG 579  
 Db 1003 PLHEAAAYAHLDLVKVFYQGERGINPAEFNEQASFPFCITIHGAPCGYSLDCDTPDRLEV 1062  
 QY 580 DAALLDAAKKGLARVKKLSLSD--NVNCRDTPQGRHSTPLHAAAGYNNLEVAEYLLQHGAD 638  
 Db 1063 VEYLSL-----KIPDINGKC-DVQ--ENTPITVAIFANKVSIILNVLVGIGAD 1106  
 QY 639 VNAQDKGGLIPLHNAASYGHVDVAALLIK-YNACVNATDKWAFPLHEAAQKRTQCLCAL 697  
 Db 1107 PNOQVYDGD-PPLYIAARQGRFEIVRCLIEVHKVDINTNRKERTALHAAARNDFMDVVKY 1165  
 QY 698 LLAHGADPTLKNQEGOTPLDLVVSADVDVSAALLTAAMPSPALPSCYKPVQLNGVRSFGATAD 757  
 Db 1166 LVROGADVNAKGIDDLRPID-IAGEKAKAYLQSS--RFLRSGHSFQ-SNEIDSEGNTHI 1220  
 QY 758 ALSSGSPSSLSAASSLDNLSCGSELSVSVSSSGTEGASLEKKEKVPGVDFSTQFVR 817  
 Db 1221 GIS-----MSARTNDKUTQOISSKGRSDSNSTEGKHSNVHVRSIDV----- 1265  
 QY 818 NLGLEHLMDFEREQITLDVLVEMGHKELKEIGINAYGHRHK-----LIKGV 864  
 Db 1266 --GALLLLDFMIRVFAS-----KKTNFAPYSGRIKTRSAQAQAALIMTERF 1311  
 QY 865 ERLISGQOGLNPYLTNTSGSGTILIDLSPPDKKEFQSVSEEMQSTVREHR 914  
 Db 1312 ENLLSGLIG-----DPIDSIDFSNVHSKIYKAIMSGR 1344

Search completed: February 12, 2003, 07:42:40  
 Job time : 27.6239 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 21:15:54 ; Search time 13.6325 Seconds  
(without alignments)  
3346.711 Million cell updates/sec

Title: US-09-843-159B-3  
Perfect score: 5769  
Sequence: 1 GFGKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGWVG 1100

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5766	99.9	1166	1	TNK2_HUMAN
2	4889.5	84.8	1327	1	TNK1_HUMAN
3	875.5	15.2	4377	1	ANK3_HUMAN
4	860.5	14.9	3924	1	ANK2_HUMAN
5	842	14.6	1880	1	ANK1_HUMAN
6	837.5	14.5	1862	1	ANK1_MOUSE
7	707.5	12.3	1059	1	Y379_HUMAN
8	571	9.9	768	1	YB23_HUMAN
9	464.5	8.1	1401	1	LATA_LATMA
10	420.5	7.3	603	1	V162_FOWPV
11	394.5	6.8	747	1	V222_FOWPV
12	390.5	6.8	832	1	ANK3_HUMAN
13	380.5	6.6	668	1	V244_FOWPV
14	368	6.4	596	1	V024_FOWPV
15	353.5	6.1	1431	1	DAPK_HUMAN
16	351	6.1	542	1	V155_FOWPV
17	350	6.1	587	1	ASB2_HUMAN
18	334.5	5.8	775	1	ASB3_MOUSE
19	327.5	5.7	525	1	V228_FOWPV
20	322.5	5.6	776	1	ANK5_HUMAN
21	322	5.6	525	1	ASB3_MOUSE
22	321	5.6	436	1	V245_FOWPV
23	319	5.5	592	1	V246_FOWPV
24	310	5.4	528	1	PPGA_HUMAN
25	309	5.4	524	1	PPGA_MOUSE
26	309	5.4	568	1	PPGB_MOUSE
27	306.5	5.3	518	1	ASB3_HUMAN
28	305	5.3	735	1	RNSA_MOUSE
29	303.5	5.3	568	1	PPGB_BOVIN
30	301.5	5.2	567	1	PPGB_HUMAN
31	299.5	5.2	692	1	ANK6_HUMAN
32	291.5	5.1	950	1	ORP1_HUMAN
33	287.5	5.0	656	1	FEMI_CABEL

34	286.5	5.0	741	1	RNSA_HUMAN
35	285	4.9	461	1	V218_FOWPV
36	285	4.9	1083	1	Y112_YEAST
37	281.5	4.9	482	1	V232_FOWPV
38	281.5	4.9	898	1	KBF2_HUMAN
39	278	4.8	984	1	KBF1_CHICK
40	277.5	4.8	429	1	AS10_HUMAN
41	277	4.8	1964	1	NTC4_MOUSE
42	276.5	4.8	434	1	V219_FOWPV
43	275	4.8	347	1	GABC_MOUSE
44	274	4.7	382	1	GABB_MOUSE
45	272.5	4.7	968	1	KBF1_HUMAN

ALIGNMENTS

RESULT 1  
TNK2\_HUMAN STANDARD; PRT: 1166 AA.  
ID TNK2\_HUMAN  
AC O9H2K2: O9HAS4: Q9H8F2:  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-  
interacting ankyrin-related ADP-ribose polymerase 2) (Tankyrase-like  
protein) (Tankyrase-related protein).  
GN TNKS2 OR TNKL OR TANK2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Fetal brain;  
RC MEDLINE=21072550; PubMed=11205898;  
RX Monz D., Munnia A., Cattesse N., Fischer U., Steudel W.-I., Feiden W.,  
RA Glass B., Meese E.U.;  
RT "Novel tankyrase-related gene detected with meningioma-specific  
sera.";  
RL Clin. Cancer Res. 7:113-119(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Breast carcinoma;  
RX MEDLINE=21190090; PubMed=11294570;  
RA Kulmov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,  
RA Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;  
RT "Cloning and characterization of TNKL, a member of tankyrase gene  
family.";  
RL Genes Immun. 2:52-55(2001).  
RN [3]  
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.  
RC TISSUE=Liver;  
RX MEDLINE=21264473; PubMed=11278563;  
RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,  
RA Eyre H.J., Sutherland G.R., Daly R.J.;  
RT "Identification of a novel human tankyrase through its interaction  
with the adaptor protein Grb14.";  
RL J. Biol. Chem. 276:17172-17180(2001).  
RN [4]  
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.  
RC TISSUE=Placenta;  
RX MEDLINE=21443728; PubMed=11454873;  
RA Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D.,  
RA Morin G.B., Yaswen P., Campisi J.;  
RT "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes  
rapid induction of cell death upon overexpression.";  
RL J. Biol. Chem. 276:35891-35899(2001).  
RN [5]  
RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND  
LNPEP/OTASE.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=21661461; PubMed=11802774;

Q05823	homo sapien
Q91517	fowlpox vir
P40480	saccharomyc
Q91503	fowlpox vir
Q00653	homo sapien
Q04861	gallus gall
Q8wx13	homo sapien
P31695	mus musculus
Q91516	fowlpox vir
Q00421	mus musculus
Q00420	mus musculus
P19838	homo sapien

RA Sbio J.I., Lodish H.F., Chi N.-W.;  
 RT "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1  
 RT (telomere-repeat-binding factor 1) and TRAP (insulin-responsive  
 RT aminopeptidase).";  
 RL Biochem. J. 361:451-459(2002).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RA Yin Y., Gelmann E.P.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RN SEQUENCE OF 838-1151 FROM N.A.  
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
 RA Tanase T., Nomura Y., Fogliya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RN FUNCTION, AND ADP-RIBOSYLATION.  
 RP MEDLINE-21602874; PubMed=11739745;  
 RX Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;  
 RA "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2  
 RT at human telomeres.";  
 RL Mol. Cell. Biol. 22:332-342(2002).  
 CC -!- FUNCTION: May regulate vesicle trafficking and modulate the  
 CC subcellular distribution of SLIC2A4/GLUT4-vesicles. Has PARP  
 CC activity and can modify TRF1, and thereby contribute to the  
 CC regulation of telomere length.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboyl](N)-acceptor =  
 CC nicotinamide + [ADP-D-riboyl](N+1)-acceptor  
 CC -!- SUBUNIT: Oligomerizes and associates with TNKS. Interacts with the  
 CC cytoplasmic domain of LNP/OTase in SLIC2A4/GLUT4-vesicles. Binds  
 CC to the N-terminus of Grb14 and TRF1 with its ankyrin repeat  
 CC region.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and  
 CC with juxtanuclear SLIC2A4/GLUT4-vesicles. Also found around the  
 CC pericentriolar matrix of mitotic centrosomes. During interphase, a  
 CC small fraction of TNKS2 is found in the nucleus, associated with  
 CC TRF1.  
 CC -!- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,  
 CC liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral  
 CC blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,  
 CC ovary, small intestine, colon, mammary gland, breast and breast  
 CC carcinoma, and in common-type meningioma. Highly expressed in  
 CC fetal liver, heart and brain.  
 CC -!- PTM: ADP-ribosylated (-auto).  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -!- SIMILARITY: CONTAINS 15 ANK REPEATS.  
 CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF305081; AAG25674.1; ALT\_INIT.  
 DR EMBL; AF264912; AAG44694.1; -  
 DR EMBL; AF329696; AAK13463.1; -  
 DR EMBL; AF342982; AAK25811.1; -  
 DR EMBL; AF309033; AAK82330.1; -  
 DR EMBL; AF438201; AAL40795.1; -  
 DR EMBL; AL359707; CAC78760.1; -  
 DR EMBL; AK023746; BAB14665.1; ALT\_INIT.  
 DR HSP; Q00420; IAWC.

DR Genew; HGNC:15677; TNKS2.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR001660; SAM.  
 DR Pfam; PF00023; ank; 20.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 15.  
 DR SMART; SM00454; SAM; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 15.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50105; SAM\_DOMAIN; 1.  
 KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.  
 FT REPEAT 57 89  
 FT REPEAT ANK 1.  
 FT REPEAT 90 122  
 FT REPEAT ANK 2.  
 FT REPEAT 123 155  
 FT REPEAT ANK 3.  
 FT REPEAT 210 242  
 FT REPEAT ANK 4.  
 FT REPEAT 243 275  
 FT REPEAT ANK 5.  
 FT REPEAT 276 308  
 FT REPEAT ANK 6.  
 FT REPEAT 363 398  
 FT REPEAT ANK 7.  
 FT REPEAT 399 431  
 FT REPEAT ANK 8.  
 FT REPEAT 432 464  
 FT REPEAT ANK 9.  
 FT REPEAT 525 557  
 FT REPEAT ANK 10.  
 FT REPEAT 591 623  
 FT REPEAT ANK 11.  
 FT REPEAT 678 710  
 FT REPEAT ANK 12.  
 FT REPEAT 711 743  
 FT REPEAT ANK 13.  
 FT REPEAT 744 776  
 FT REPEAT ANK 14.  
 FT REPEAT ANK 15.  
 FT DOMAIN 873 936 SAM.  
 FT DOMAIN 1023 1162 PARP.  
 FT CONFLICT 331 337 KGHSLQ -> QRPLVA (IN REF. 1).  
 FT CONFLICT 357 361 NFRHP -> IQAS (IN REF. 1).  
 FT CONFLICT 966 966 Q -> P (IN REF. 8).  
 SQ SEQUENCE 1166 AA; 126917 MW; 4C8B3B8D97CE704 CRC64;  
 Query Match 99.9%; Score 5766; DB 1; Length 1166;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GFGKRDVVEYLLQNGASVQARDGGLIPLHNACSFHAEVYVNVLLLRHGADPNARDNNYT 60  
 DB 67 GFGKRDVVEYLLQNGANVOARDGGLIPLHNACSFHAEVYVNVLLLRHGADPNARDNNYT 126  
 QY 61 PLHEAAIKGIDVCIVLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKDELLSAR 120  
 DB 127 PLHEAAIKGIDVCIVLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKDELLSAR 186  
 QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 180  
 DB 187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLV 246  
 QY 181 PLHNACSYGHVEYTELLVKGACVNMALWQFTPLHFAASKNRVEYCSLLSYGADPTLL 240  
 DB 247 PLHNACSYGHVEYTELLVKGACVNMALWQFTPLHFAASKNRVEYCSLLSYGADPTLL 306  
 QY 241 NCHNKAIDLAPTPQLKERLAYEFKHSLLQAAAREADVTIRKKHLSLEMYNFKHPQTHET 300  
 DB 307 NCHNKAIDLAPTPQLKERLAYEFKHSLLQAAAREADVTIRKKHLSLEMYNFKHPQTHET 366  
 QY 301 ALHCAASPYPKRKQICEILLRKGANINEKTEFLTPHVASEKAHNDVVEYVVKHEAKV 360  
 DB 367 ALHCAASPYPKRKQICEILLRKGANINEKTEFLTPHVASEKAHNDVVEYVVKHEAKV 426  
 QY 361 NALDNLGQTSLHRAAYCGHLOTCRLLSYCCDPNIIISLQGTALQMGNNVQQLLEGIS 420  
 DB 427 NALDNLGQTSLHRAAYCGHLOTCRLLSYCCDPNIIISLQGTALQMGNNVQQLLEGIS 486  
 QY 421 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGQSTPLHFAAGYNRVSVVEYLL 480  
 DB 487 LGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGQSTPLHFAAGYNRVSVVEYLL 546  
 QY 481 QHGADVHAKDKGLVPLHNACSYGHVEYAEVLVKGAVNVADLWKFTEPLHEAAKGYE 540

```

Db 547 QHGVADVHAKDGLVPLHNACSYGHVEAELLVKGAVVNVADLWKFPLHEAAAKGYE 606
Qy 541 ICKLLLOHGADPTKKNRGNTPLDLVKDGDITDIQDLLRCDAAALLDAAKGGLARVKKLSS 600
Db 607 ICKLLLOHGADPTKKNRGNTPLDLVKDGDITDIQDLLRCDAAALLDAAKGGLARVKKLSS 666
Qy 601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLPLHNAASYGHVD 660
Db 667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLPLHNAASYGHVD 726
Qy 661 VAALLIKYNACVNAQDKWAFPLPHEAAQGRQTLQCALLLAHGAADPTLKNQEGQTPLDLVS 720
Db 727 VAALLIKYNACVNAQDKWAFPLPHEAAQGRQTLQCALLLAHGAADPTLKNQEGQTPLDLVS 786
Qy 721 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLSG 780
Db 787 ADDVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGSPSSLSAASSLDNLSG 846
Qy 781 SFSELSVVSSSGTEGASLEKKEVPGVDFSTQFVRNGLGLEHMDIPEREQITLDVLVE 840
Db 847 SFSELSVVSSSGTEGASLEKKEVPGVDFSTQFVRNGLGLEHMDIPEREQITLDVLVE 906
Qy 841 MGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSOTILDLSPDDKEFQ 900
Db 907 MGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSOTILDLSPDDKEFQ 966
Qy 901 SYVEEMOSTVREHROGGHAGGIFNRYNLIKIOKCNKLLWERYTHRRKEVSEENHNHANE 960
Db 967 SYVEEMOSTVREHROGGHAGGIFNRYNLIKIOKCNKLLWERYTHRRKEVSEENHNHANE 1026
Qy 961 RMLFHGSPFVNIAIHKGFDERHAYIGGMFAGIYFAENSCKSNQYVYIGGTCPCPVHKD 1020
Db 1027 RMLFHGSPFVNIAIHKGFDERHAYIGGMFAGIYFAENSCKSNQYVYIGGTCPCPVHKD 1086
Qy 1021 RSCYCHROLLFCRVTLGKSFLOFSAMKMAHSPGHSHSVTGRPSVNGLAALAEYVIYRGQ 1080
Db 1087 RSCYCHROLLFCRVTLGKSFLOFSAMKMAHSPGHSHSVTGRPSVNGLAALAEYVIYRGQ 1146
Qy 1081 AYPEYLITYOIMRPEGMDG 1100
Db 1147 AYPEYLITYOIMRPEGMDG 1166

RESULT 2
TNK1 HUMAN
ID TNK1 HUMAN STANDARD; PRT; 1327 AA.
AC O95271; O95272;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase 1) (TNKS-1) (TRF1-
interacting ankyrin-related ADP-ribose polymerase).
GN TNKS OR TNKS1 OR TINI OR TINFL OR PARPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Testis;
RX MEDLINE=99040105; PubMed=9822378;
RA Smith S., Giriat I., Schmitt A., de Lange T.;
RT "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
RL Science 282:1484-1487(1998).
RN [2]
RP SUBCELLULAR LOCALIZATION.
RX MEDLINE=99454782; PubMed=10523501;
RA Smith S., de Lange T.;
RT "Cell cycle dependent localization of the telomeric PARP, tankyrase,
to nuclear pore complexes and centrosomes.";
RL J. Cell Sci. 112:3649-3656(1999).
RN [3]
RP FUNCTION, AND PHOSPHORYLATION.
RX MEDLINE=20556282; PubMed=10988299;
Chi N.-W., Lodish H.F.;
RT "Tankyrase is a golgi-associated mitogen-activated protein kinase
substrate that interacts with IRAP in GLUT4 vesicles.";
RN J. Biol. Chem. 275:38437-38444 (2000).
[4]
FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
RX MEDLINE=21602874; PubMed=11739745;
Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
at human telomeres.";
RL Mol. Cell. Biol. 22:332-342(2002).
CC !- FUNCTION: May regulate vesicle trafficking and modulate the
subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
activity and can modify TRF1, and thereby contribute to the
regulation of telomere length.
CC !- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor =
nicotinamide + [ADP-D-riboseyl](N)-acceptor.
CC !- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
the cytoplasmic domain of INPEP/Otase in SLC2A4/GLUT4-vesicles.
CC Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC !- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is
also found at nuclear pore complexes and around the pericentriolar
matrix of mitotic centrosomes. During interphase, a small fraction
of TNKS is found in the nucleus, associated with TRF1.
CC !- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
produced by alternative splicing.
CC !- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC !- PTM: Upon insulin-stimulation, phosphorylated on serine residues
by MAPK kinases.
CC !- PTM: ADP-ribosylated (-auto).
CC !- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC !- SIMILARITY: CONTAINS 15 ANK REPEATS.
CC !- SIMILARITY: CONTAINS 1 SAM DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF082556; AAC79841.1; -
DR EMBL; AF082557; AAC79842.1; -
DR EMBL; AF082558; AAC79843.1; -
DR EMBL; AF082559; AAC79844.1; -
DR HSSP; Q00420; IAWC.
DR Genew; HGNC:11941; TNKS.
DR MIM; 603303; -
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 21.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 15.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 15.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
KW Phosphorylation; Alternative splicing.
FT REPEAT 215 247 ANK 1.
FT REPEAT 248 280 ANK 2.
FT REPEAT 281 313 ANK 3.
FT REPEAT 368 400 ANK 4.
FT REPEAT 401 433 ANK 5.
FT REPEAT 434 466 ANK 6.
FT REPEAT 521 556 ANK 7.
FT REPEAT 557 589 ANK 8.
FT REPEAT 590 622 ANK 9.

```

```
FT REPEAT 683 715 ANK 10.
FT REPEAT 716 748 ANK 11.
FT REPEAT 749 781 ANK 12.
FT REPEAT 836 868 ANK 13.
FT REPEAT 869 901 ANK 14.
FT REPEAT 902 934 ANK 15.
FT DOMAIN 1030 1089 SAM.
FT DOMAIN 1176 1327 PARP.
FT DOMAIN 9 14 POLY-HIS.
FT DOMAIN 27 34 POLY-PRO.
FT DOMAIN 128 134 POLY-SER.
FT DOMAIN 137 145 POLY-SER.
FT VARSPLIC 641 643 EST -> GHS (IN ISOFORM 2).
FT VARSPLIC 644 1327 MISSING (IN ISOFORM 2).
FT MUTAGEN 1184 1184 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT MUTAGEN 1291 1291 WITH A-1291.
FT MUTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT SEQUENCE 1327 AA: 142010 MW: 144DE985C710B957 CRC64;
Query Match 84.8%; Score 4889.5; DB 1; Length 1327;
Best Local Similarity 83.0%; Pred. No. 3-2e-292;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;
QY 1 GFGKDVVEYLLQNGASVOARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYNT 60
DB 225 GFGKDVVEYLLQNGASVOARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYNT 284
QY 61 PLHEAAIKGKIDVICVILLOHGAAPTIRNTDGTALDLDADPSAKAVLTGEYKDDLEESAR 120
DB 285 PLHEAAIKGKIDVICVILLOHGAAPTIRNTDGTALDLDADPSAKAVLTGEYKDDLEESAR 344
QY 121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLOHGAADVHAKDGLV 180
DB 345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLOHGAADVHAKDGLV 404
QY 181 PLHNACSYGHEYTELLVHKGACVNDMLWQFTPLHEAASKNREVCVSLLSYGADPTLL 240
DB 405 PLHNACSYGHEYTELLVHKGACVNDMLWQFTPLHEAASKNREVCVSLLSYGADPTLL 464
QY 241 NCHNSAIDLAPTPOLKERLAYEFKSHLSLQAAREADVTIRIKKHSLEVMVNFKHPOTHET 300
DB 465 NCHGKSAMDAPTPOLKERLAYEFKSHLSLQAAREADVTIRIKKHSLEVMVNFKHPOTHET 524
QY 301 ALHCAASAPYPRKQTCCELLLRKANINKEFTPLHVASAKHNDVVEVYVVEVYVVEVY 360
DB 525 ALHCAVASLHPRKQTCCELLLRKANINKEFTPLHVASAKHNDVVEVYVVEVYVVEVY 584
QY 361 NALDNLGQTSLHRAAYCGHLCQRLALLSYGCDPNITSLQGTALONGNENVOQLLOEGIS 420
DB 585 NALDNLGQTSLHRAAYCGHLCQRLALLSYGCDPNITSLQGTALONGNENVOQLLOEGIS 644
QY 421 LGNSEADROLLEAAKAGDVETVKCLTQVSNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
DB 645 IRTSDVDYRLLEASKAGDLETKQLCSSQNVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 704
QY 481 QHGADVHAKDKGGLVPLHNACSYGHEVEAEVLLVKGAVNVNADLWFTPLHEAAAKGKYE 540
DB 705 QHGADVHAKDKGGLVPLHNACSYGHEVEAEVLLVKGAVNVNADLWFTPLHEAAAKGKYE 764
QY 541 ICKLLLOHGAADPTKKNRDCNTPLDLVKDGTDTQDLRLGDAALDAKKGCLARVKLLSS 600
DB 765 ICKLLLOHGAADPTKKNRDCNTPLDLVKDGTDTQDLRLGDAALDAKKGCLARVKLLSS 824
QY 601 PDNVNCRDTQGRHSTPLHLAAGYNLLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 660
DB 825 PDNVNCRDTQGRHSTPLHLAAGYNLLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVD 884
QY 661 VAALLIKYNACVNDTKWFTPLHEAAQGRQTCALLAHGADPTLNQEGTPTLDLVS 720
DB 885 VAALLIKYNACVNDTKWFTPLHEAAQGRQTCALLAHGADPTLNQEGTPTLDLVS 944
QY 721 ADDVSALLTAAMPSPALPSCYKQVINGVRSPGATADALSSGSPSSLSAASSLDNLSSG 780
```

```
DB 945 ADDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTPSCLSAASIDNLITG 997
QY 781 SFSELSVSVSSSGTEGASSLEKK--EVPGVDFSTQVFNRLGLEHLMDFEREQITLDVL 838
DB 998 PLAEAVGGASNAGDGAAGTERKEGEVAGLDMNISQPLKSLGLEHLRDFETEQITLDVL 1057
QY 839 VEMGHKELKEIGINAYCHRHKLKGVVERLISGOGLNPYLTLNTISGSGTILIDLSPDKE 898
DB 1058 ADMGHHELKEIGINAYCHRHKLKGVVERLISGOGLNPYLTLNTISGSGTILIDLSPDKE 1117
QY 899 FQSVSEEMQSTVREHRDGGHAGIFNRYNLIKQVKCNKKILWERYTHRRKEVSEENHHA 958
DB 1118 YQSVSEEMQSTVREHRDGGHAGIFNRYNLIKQVKCNKKILWERYTHRRKEVSEENHHA 1177
QY 959 NERMLFHGSPFVNAIHKGFDERHAYIGMGFAGIYFAENSCKSNQYVYIGGGTGCPTH 1018
DB 1178 NERMLFHGSPFVNAIHKGFDERHAYIGMGFAGIYFAENSCKSNQYVYIGGGTGCPTH 1237
QY 1019 KDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGLALAEYVIYRG 1078
DB 1238 KDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGLALAEYVIYRG 1297
QY 1079 EQAYPEYLITYQIMRPE 1095
DB 1298 EQAYPEYLITYQIMRPE 1314
RESULT 3
ANK3_HUMAN STANDARD; PRT: 4377 AA.
ID ANK3_HUMAN Q12955;
AC Q12955;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RT at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
CC -!- FUNCTION: Membrane-cytoskeleton linker.
CC -!- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
CC alternative splicing.
CC -!- TISSUE SPECIFICITY: Expressed in brain and other tissues.
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U13616; AAA64834.1; -.
DR HSP; P55273; 1B78.
DR Genew; HGNC:494; ANK3.
DR MIM; 600465; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
```

```

DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102 ANK 1.
FT REPEAT 106 135 ANK 2.
FT REPEAT 139 168 ANK 3.
FT REPEAT 172 201 ANK 4.
FT REPEAT 203 230 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 267 296 ANK 7.
FT REPEAT 300 329 ANK 8.
FT REPEAT 333 362 ANK 9.
FT REPEAT 366 395 ANK 10.
FT REPEAT 399 428 ANK 11.
FT REPEAT 432 461 ANK 12.
FT REPEAT 465 494 ANK 13.
FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.
FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT DOMAIN 1519 1898 SER-RICH.
FT DOMAIN 4090 4174 DEATH.
SQ SEQUENCE 4377 AA; 480399 MW; F42379E5576B8684 CRC64;

Query Match 15.2%; Score 875.5; DB 1; Length 4377;
Best Local Similarity 30.5%; Pred. No. 4.4e-45;
Matches 273; Conservative 126; Mismatches 362; Indels 135; Gaps 21;

QY 3 GRKDVVEYLLQNGASVQARDGGGLIPLHNACSGHAEVNVLLLRHGADPNARDNWTPL 62
DB 85 GHVEVSELLQREANVDAAATKNGNTALHIASLAGQAEVVKVLYVTNGANVNAQSQNGTPL 144
QY 63 HEAAIKGIDKIVLLQHGAEPTIRNTDGR TALDA-----DPSAKAVLGEYK---KDE 114
DB 145 YMAAOEHLEVVKFLDNGASQSIEDGTPTPLAVALQQHDQVSVLLENDTKGKVRLP 204
QY 115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLPLAAGYNRYKIVQLLIQHGADVHAK 174
DB 205 ALHIAARKDDTKAAALLQNDNNADVESKSGTTPHIAAHYGNINVTALLNRAAAVDFT 264
QY 175 DKGLVPLHNACSGHYEVTELLVKGACVNMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
DB 265 ARNDITPLHVASKRGANVMVKLLDRGAKTDKTRDGLTPLHCGARSQHGQVVEMLLDR 324
QY 235 ADPTLLCNKNSAIDAPTLPOLKERLAYEFKGSLLQAAAREADVTRIKKHLSEWVFKH 294
DB 325 A-PLSKTKN-----GLSPLHATQGD-----HLNCVQLLIQH 356
QY 295 -----POTHETALHCAASAPYKPKQICELLRKGANINEKTKFTPLPLHVASERAKHN 347
DB 357 NVPVDDTNDYLTALHVAACHGHYK---VAKVLLDKKANPAKALNGFTPLHIACKKNRI 413
QY 348 DVVEVVKHAKVNALDNLGTSLHRAAYCCHLOTCLRLLSYCGDPNIIISLQGTALQMG 407
DB 414 KVMELLKHGASIQAVTESGLTPTTHVAFMGHVIVISQLMHHGASPNNTNVRGETALHMA 473
QY 408 -----NENVOOLLQEGISLGNSEADRO--LLEAAKAGDVETVKKLCTV-QSVNCRDIEGR 459
DB 474 ARSQAEVVRVLYQDGAQVENKAKDDTPTPLHISARLKGADIVQQLQGGASPNATTSG- 532
QY 460 QSTPLHFAAGYNRSVVEYLLQHGADVHAKDKGGLVPLHNACSGHYEVAELLVKHGAVV 519

```

CC -1- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL  
 CC CELLS THROUGHOUT THE BRAIN.  
 CC -1- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES  
 CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE  
 CC AND FUNCTION (POTENTIAL).  
 CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.  
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X56957; CAA40278.1; -;  
 CC EMBL; X56958; CAA40279.2; -;  
 CC EMBL; Z26634; CAA42644.1; -;  
 CC EMBL; M37123; AAG62828.1; -;  
 CC PIR; S14533; S14533.  
 CC PIR; A39643; A39643.  
 CC PIR; B39643; B39643.  
 CC PIR; S14569; S14569.  
 CC HSSP; P42771; IDC2.  
 CC Genew; HGNC:493; ANK2.  
 CC MIM; 106410; -;  
 CC InterPro; IPR002110; ANK.  
 CC InterPro; IPR000488; Death.  
 CC InterPro; IPR000906; ZU5.  
 CC Pfam; PF00023; ank; 24.  
 CC Pfam; PF00531; death; 1.  
 CC Pfam; PF00791; ZU5; 1.  
 CC PRINTS; PR01415; ANKYRIN.  
 CC SMART; SM00248; ANK; 21.  
 CC SMART; SM00005; DEATH; 1.  
 CC SMART; SM00218; ZU5; 1.  
 CC PROSITE; PS50088; ANK\_REPEAT; 20.  
 CC PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 CC PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;  
 KW Phosphorylation.  
 FT REPEAT 63 92 ANK 1.  
 FT REPEAT 96 125 ANK 2.  
 FT REPEAT 129 158 ANK 3.  
 FT REPEAT 162 191 ANK 4.  
 FT REPEAT 193 220 ANK 5.  
 FT REPEAT 232 261 ANK 6.  
 FT REPEAT 265 294 ANK 7.  
 FT REPEAT 298 327 ANK 8.  
 FT REPEAT 331 360 ANK 9.  
 FT REPEAT 364 393 ANK 10.  
 FT REPEAT 397 426 ANK 11.  
 FT REPEAT 430 459 ANK 12.  
 FT REPEAT 463 492 ANK 13.  
 FT REPEAT 496 525 ANK 14.  
 FT REPEAT 529 558 ANK 15.  
 FT REPEAT 562 591 ANK 16.  
 FT REPEAT 595 624 ANK 17.  
 FT REPEAT 628 657 ANK 18.  
 FT REPEAT 661 690 ANK 19.  
 FT REPEAT 694 723 ANK 20.  
 FT REPEAT 727 756 ANK 21.  
 FT REPEAT 760 789 ANK 22.  
 FT REPEAT 793 822 ANK 23.  
 FT DOMAIN 1773 1950 REPEAT-RICH REGION.  
 FT REPEAT 1773 1784 REPEAT A.  
 FT REPEAT 1785 1796 REPEAT A.  
 FT REPEAT 1797 1808 REPEAT A.  
 FT REPEAT 1809 1820 REPEAT A.  
 FT REPEAT 1821 1832 REPEAT A.  
 FT REPEAT 1833 1844 REPEAT A.  
 FT REPEAT 1845 1856 REPEAT A.

FT REPEAT 1857 1867 REPEAT A (APPROXIMATE).  
 FT REPEAT 1868 1879 REPEAT A.  
 FT REPEAT 1880 1891 REPEAT A.  
 FT REPEAT 1892 1902 REPEAT A (APPROXIMATE).  
 FT REPEAT 1903 1914 REPEAT A.  
 FT REPEAT 1915 1926 REPEAT A.  
 FT REPEAT 1927 1938 REPEAT A.  
 FT REPEAT 1939 1950 REPEAT A.  
 FT DOMAIN 3536 3620 DEATH.  
 FT VARSPIC 1039 1039 Q -> QFLGLHLPTAPPPLNEGESLSVRIQLQGPQTK  
 FT VARSPIC 1444 3528 (IN ISOFORM 2).  
 FT CONFLICT 475 476 MISSING (IN ISOFORM 2 AND ISOFORM 3).  
 FT CONFLICT 971 971 GQ -> PE (IN REF. 4).  
 FT CONFLICT 3581 3582 I -> S (IN REF. 1).  
 FT CONFLICT 3586 3586 QY -> HA (IN REF. 1).  
 FT CONFLICT 3586 3586 I -> Y (IN REF. 1).  
 SQ SEQUENCE 3924 AA: 430337 MW: 52AC496C428E29D2 CRC64;  
 Query Match 14.9%; Score 860.5; DB 1; Length 3924;  
 Best Local Similarity 29.1%; Pred. No. 3.1e-44;  
 Matches 297; Conservative 147; Mismatches 404; Indels 174; Gaps 29;  
 QY 3 GRKDVVEYLLQNGASVOARDGGLIPLHNACSGHAEVNVNLLRHGADPNARDNWNVTPL 62  
 DB 75 GHVGLVQELLGRGSSVDSATKGTALHSLAGQAQAEVVKLVKREGANINQAQSONGFTPL 134  
 QY 63 HEAAIKGIDVCIVLLQHGAEPTIRNTDGRT---ALDLADPSAKAVLTGCKKDEL--- 115  
 DB 135 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQOQHNAVAILLNDTKGKVRLP 194  
 QY 116 -LESARSGNEEKWMALLTPLNVNCHASDGRKS-----TPLHAAAGYNRKIVQL 163  
 DB 195 ALHIAARKODTKSAALLQ---NDHNADVQSKMNVNRTTESGFTPLHIAAHYGNVAVTL 251  
 QY 164 LLOHGADVHAKDKGLVPLHNACSYGHYEYVELLVKHGACVNDMLWQFTPLHEAASKNR 223  
 DB 252 LLNRGAADVFTARGITPLHVASKRGNTNMVKLLDRGGQIDAKTRDGLTPLCAARSGH 311  
 QY 224 VEVCSSLLSYGADPTLLNCHNKSALDLPQPKERLAYEFKSHSLQAAAREADVTRIKK 283  
 DB 312 DOVVELLERGA-PLLARTKN-----GLSPLHMAAQGDHVECVK 349  
 QY 284 HLSLEMVNFKHP-----QTHETALHCAASAPYKPKQICELLRLKRGANINEKTEFTPL 338  
 DB 350 HL----LQHKAPVDDVTLYLTALHVA---HCGHYRVTKLLDKRANPNARALNGFTPL 402  
 QY 339 HVASERAHNDVVEVVKHAKVNALDNLGOTSLHRAAYCGHLOTCLLLSYGCDPNILSL 398  
 DB 403 HIACKNRKVMELLVKYGIASIOAITESGLTPIHVAAFMGHNLNIVLLLLQNGASPDVTNI 462  
 QY 399 QGFALQMGNNVQQLLOEGISLGNSEADROLLEAAKAGDVETVKKICTVQSVNCRDIEG 458  
 DB 463 RGETALHM-----AARAGQEVVR--CLLRNGALVDARA 494  
 QY 459 R-QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVELLVKHA 517  
 DB 495 REEQTPHIAASRLGKTEIVQLLQHMAHPDAATTNGYTPHLHISAREQVDVASVLEAGA 554  
 QY 518 VNVADLWKTPLHEAAAKGYEICKLLQHGADPTKKNRDGNTPLDLVKDGDGTDIDQLL 577  
 DB 555 AHSLATKKGFTPLHVAAKYGSGLDVAKLLQRRRAAASAGKNGLTPLHVAAHYDN----- 608  
 QY 578 RGDALLDAAKKGLARVKKLSPPNVNCRDQTGRHSTPLHAAAGYNNLEVAEYLLQHA 637  
 DB 609 -----QKVALLLEKGA---HATAKNGYTPHLHIAAKKNQMQIASTLLNYGA 653  
 QY 638 DVNAQDKGLIPLHNAASYGHVDVAALLTKYNACVNATKWAFTPLHEAAQKGTOLCAL 697  
 DB 654 ETNIVTKOGVTPHLASQEGHTDMVTLLLDKGANIHMTKSGLTSLHAAQEDKVNVAI 713  
 QY 698 LLAHGADPTLKNOEGOTPLDLV----SADDVSALLTAAMPSPA-LPSCYKP----- 743  
 DB 714 LTKHGADQDAHTKLGYPPLIVACHYGNVKNVNFLLKQGANVNAKTKNGYTPHQAQAQGH 773





```
FT VARIANT 20 20 R -> T.
FT VARIANT 462 462 /FTid=VAR_000595.
FT VARIANT 618 618 V -> I (IN HS).
FT VARIANT 618 618 /FTid=VAR_000596.
FT VARIANT 749 749 R -> H (IN BRUEGEN).
FT VARIANT 749 749 /FTid=VAR_000597.
FT VARIANT 844 844 V -> A.
FT VARIANT 844 844 /FTid=VAR_000598.
FT VARIANT 1285 1285 D -> E.
FT VARIANT 1285 1285 /FTid=VAR_000599.
FT VARIANT 1391 1391 E -> D.
FT VARIANT 1391 1391 /FTid=VAR_000601.
FT VARIANT 1591 1591 S -> T.
FT VARIANT 1591 1591 /FTid=VAR_000602.
FT VARIANT 1698 1698 D -> N (IN DESSELDORF).
FT VARIANT 1698 1698 R -> D.
FT CONFLICT 229 229 /FTid=VAR_000603.
FT CONFLICT 1545 1545 A -> S (IN REF. 2).
FT CONFLICT 1545 1545 V -> I (IN REF. 2).
SQ SEQUENCE 1880 AA; 206145 MW; 1C5F5E7EFD1CD428 CRC64;

Query Match 14.6%; Score 842; DB 1; Length 1880;
Best Local Similarity 29.9%; Pred. No. 1.4e-43;
Matches 240; Conservative 122; Mismatches 264; Indels 177; Gaps 20;

QY 3 GRKDVVEYLQNGASVOARDGGLPLHNACSGFHAQVNVNLLRHGADPNARONWNTPL 62
Db 88 QODEVRELNYGANNVNAQSQKGFPLMAQENHLEVKFLENGANQNVATEDGFTPL 147
QY 63 HEAAIKGIDVCIVLQHGAEPTI-----RNTDGRITALDLPFSAKAVLTGEYK 113
Db 148 AVAQQGHENVAHLINYGTKGVRLPALHIAARNDDRTA-----AVLLQNDPNP 198
QY 114 ELLESARGNEERKMALLTPLNVNCHASDGRKS-----TPHLAAGY 155
Db 199 DVL--SKTG-----FTPLHAAHYENINVAQLLLNREGASVNFPTPONGITPLHIA 248
QY 156 NRKVIYVOLLQHGADYHAKDGDLPLHNACSGHYEVTELLVKHGACVNMALDQFTPL 215
Db 249 GNVIMVRLLDRCQAETKTKDELTPHCAARNGHVRISEILLDHGAPQAOKTKNLSPI 308
QY 216 HEAAKNRVEVCSLLLSYGADPTLLNCHNKSADLAPTOLKRLAYEFKGSLLQAARE 275
Db 309 HMAAQGDHLDCVRLLOYDAE-----ID----- 331
QY 276 ADVTRIKKHLSELMVNFKHPQTHETALHCAAAASPYPKQKQICELLRLKGANINEKTE 335
Db 332 -DIT-----LDHLTPHVA-----HCHHRVAKVLLDKGAKPNRSLNGF 371
QY 336 TPLHVAEKAHNDVVEVYVYKHEAKVNALDNGTSLHRAAYCGHGLQTCRLLSYGCDPNI 395
Db 372 TPLHIACKKNHVRVMBELLKTKGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 431
QY 396 ISLQGFALQMG-----NENVQOLQF-----GISLNSRADROLLE- 432
Db 432 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLEN 491
QY 433 -----AAKAGDVETVKKLTQV--SVNCRDTEGROSTPLHFAAGVNRYS 474
Db 492 NANPNLATTAGHTPLHIAAREGHVETVLLALLEKEASQACMTKKG--FTPLHVAAYKGV 549
QY 475 VVEYLQHGADYHAKDGGLPLHNACSGHYEVTELLVKHGAVNVADLWK--FTPLHBA 533
Db 550 VAEILLERDAHPNAAGKNGLTPLHVAVHNHNLIDVKLLPRGGSPLH--SPAWNGYTPH 608
QY 534 AAKGYEICKLLOHGADPTKKNRQGNTPDL--VKDGTDDIQLRGDALLDAKKGCL 592
Db 609 AKQNVARSLSQYGGGSAANESVQGVTPPLHAAQEGHAEWALL----- 653
QY 593 ARVKKLSPPDNYNCRDQGRHSTPLHAAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHN 652
Db 654 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKGVNVDTATRMGYTP 706
```

```
QY 653 AASYGHVDVAALLIKYNACVNATDKWAFPLHEAAQKRTQLCALLLAHAGADPTLKNOEG 712
Db 707 ASHYGNIKLVKFLQHQADVNAKTKLGYSPLHQAAQOQHTDIVTLLKNGASPNVSSDG 766
QY 713 QTPLDL-----VSADDVSALLT 729
Db 767 TPLAIARIGYISVTDVLKVV 789

RESULT 6
ID ANKL_MOUSE STANDARD; PRT; 1862 AA.
AC Q02357;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin).
GN ANKL OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte;
MEDLINE=92345717; PubMed=1386265;
RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT regulatory domain.";
RL Mamm. Genome 3:281-285(1992).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
CC -!- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).
CC -!- PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84756; AAA37236.1; -
DR HSSP; Q00420; 1AWC.
DR MGD; MGI:88024; Ank1.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
FT DOMAIN 1 827 89 KDA DOMAIN (ANION EXCHANGE PROTEIN
FT BINDING DOMAIN).
FT DOMAIN 828 1386 62 KDA DOMAIN (SPECTRIN BINDING
```

FT DOMAIN 1387 1862  
FT 55 KDA REGULATORY DOMAIN (REGULATES  
FT THE BINDING OF ANKYRIN TO SPECTRIN  
FT AND THE BAND 3 PROTEIN).  
FT ANK 1.  
FT REPEAT 40 69  
FT REPEAT 73 102  
FT REPEAT 106 135  
FT REPEAT 139 168  
FT REPEAT 170 197  
FT REPEAT 201 230  
FT REPEAT 234 263  
FT REPEAT 267 296  
FT REPEAT 300 329  
FT REPEAT 333 362  
FT REPEAT 366 395  
FT REPEAT 399 428  
FT REPEAT 432 461  
FT REPEAT 465 494  
FT REPEAT 498 527  
FT REPEAT 531 560  
FT REPEAT 564 593  
FT REPEAT 597 626  
FT REPEAT 630 659  
FT REPEAT 663 692  
FT REPEAT 696 725  
FT REPEAT 729 758  
FT REPEAT 762 791  
FT DOMAIN 1399 1483  
FT DEATH.  
FT SEQUENCE 1862 AA; 204242 MW; AE6B85B5B29001E5 CRC64;  
Query Match 14.5%; Score 837.5; DB 1; Length 1862;  
Best Local Similarity 29.5%; Pred. No. 2.7e-43;  
Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;

QY 3 GRKDVVEYLQNGASQVARDGGLPLHNACSGFGEVNVNLLRHGADPNARDNNYTP 62  
DB 85 GODEVRELVNGANYNAQSKGFTPLYMAAQENHLEVYKFLLENGANQVATEDGFTPL 144  
QY 63 HEAAIKGKIDCVTLVQHGAEPTI-----RNTDGRALDLDPSAKAVLTGEYK 113  
DB 145 AVAQQGHEVNVVAILNYGTGKVRPLPALHIAARNDDRTIA-----AVLLQNDPNP 195  
QY 114 ELLESARSGNEERKMALLTPLNYNCHASDGKRS-----TPLHLAAGY 155  
DB 196 DVL--SKTG-----FTPLHIAAHYENLVNVAQLLNRGASVNFPPQNGITPLHIASRR 245  
QY 156 NRKIVQLLQHGADYHAKDGLDPLHNACSGYHVEYVELLVKHCACVNMDLQFTPL 215  
DB 246 GNVMVRLLLDRGAQITRTKDELTPHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 305  
QY 216 HEAAKNRVVEVCSLLSYGADPTLLNCHNKSATDLAPTQPKERLAYEFKGHSLQAARE 275  
DB 306 HMAAQGDHLCVRLLLQYNAE-----ID----- 328  
QY 276 ADVTRIKKHLSEVMNFKHPQTHETALHCAASPYPKRKOICELLRKGANINEKTKEFL 335  
DB 329 -DIT-----LDHLLPLHVA-----HCHHVRVAKVLVDKGAKPNSRALNGF 368  
QY 336 TPLHVASEKAHNDVVEVYVKEAVKVNALDNGTSLHRAAYCGHLQTCRLLLSYGCDPNI 395  
DB 369 TPLHIACKKHHIRVMEILLTGTASIDAVTESGLTPHVASFMGHPLDIVKNLQRGASPNV 428  
QY 396 ISLQGFALQW-----GNENYQQLQEGISLGNSEAD-----ROLLE- 432  
DB 429 SNVKVETPLHMAARAGHTEVAKYLLONKAKANAKAKDDQTPHCAARIGHTGMVKLLLEN 488  
QY 433 -----AAKAGDVETVKICTVQ--SVNCRDIEGROSTPLHFAAGYNRVS 474  
DB 489 GASPNLATTAGHTPLHTAAREGHVDVTDALALLEKEASQACMTKKG--FTPLHVAAYGKVR 546  
QY 475 VVEYLLQHGADYHAKDGLVPLHNACSGYHVEYVELLVKHCACVNMVADLWK--FTPLHEA 533  
DB 547 LAELLLEHDAHNAAGKNGLTPLHVAHVHNNNDIVKLLPRGGSPH--SPAWNGYTPPLHTA 605

QY 534 AAKKYETICKLLQHGADPTKKNRDGNTPDL-VKDGDTDIODLLRGAALLDAKKGCL 592  
DB 606 AKQKQIEVARSLQYGGSAEVSOGVTPLHLAAQEGHTEWVALL----- 650  
QY 593 ARVKKLSPDNVNCRTDQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHN 652  
DB 651 -----LSKQANGNLGNKSG--LTPLHLVVSQGHVLVADVLKIKHGVTDATTRMGYTP 703  
QY 653 AASGHVDVVAALLIKYNACVNAWKWFTPLHAAQKGRTOLCALLAHAGADPTLKNOEG 712  
DB 704 ASHYGNIKLVKFLQHQADVNAKTKLGYSPHQAQOQHTDIVTLLKNGASNEPVSNG 763  
QY 713 QTPLDL-----VSADDVSALLT 729  
DB 764 TTPLAIARLGYISVTDVLKVV 786  
RESULT 7  
Y379\_HUMAN STANDARD; PRT; 1059 AA.  
ID Y379\_HUMAN STANDARD; PRT; 1059 AA.  
AC O15084;  
DC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein KIAA0379 (Fragment).  
GN KIAA0379.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=97349984; PubMed=9205841;  
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. VII.  
RT The complete sequences of 100 new cDNA clones from brain which can  
RT code for large proteins in vitro."  
RL DNA Res. 4:141-150(1997).  
CC -!- SIMILARITY: CONTAINS AT LEAST 27 ANK REPEATS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB002377; BAA20833.2; -  
CC HSP; P80144; 2MYO.  
CC InterPro; IPR002110; ANK.  
CC Pfam; PF00023; ank; 28.  
CC SMART; SM00248; ANK; 26.  
CC PROSITE; PS50088; ANK\_REPEAT; 24.  
CC PROSITE; PS50297; ANK\_REPEAT; 1.  
CC Hypothetical protein; Repeat; ANK repeat.  
CC NON\_TER 1 1  
DR REPEAT 46 75 ANK 1.  
DR REPEAT 79 108 ANK 2.  
DR REPEAT 112 141 ANK 3.  
DR REPEAT 145 174 ANK 4.  
DR REPEAT 178 207 ANK 5.  
DR REPEAT 211 240 ANK 6.  
DR REPEAT 244 273 ANK 7.  
DR REPEAT 277 307 ANK 8.  
DR REPEAT 311 340 ANK 9.  
DR REPEAT 344 373 ANK 10.  
DR REPEAT 377 406 ANK 11.  
DR REPEAT 410 439 ANK 12.  
DR REPEAT 443 472 ANK 13.



```
QY 44 LLRHGADPNARDNNWNTYPLHEAAATGKIDVCIVLLQHGAEPTIRNTDGRATDLADPSAK 103
DB 2 LLQGAENVKADNEGRTALIAAAYNGHREIVEHLLDHGAENVHEDVDGRTALSA-----56
QY 104 AVLTGEYKKDELLESARSNGEKKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOL 163
DB 57 -----ALCVP-----ASKGHAS-----VVSL 72
QY 164 LLOHGADYHAKDGLVPLHNACSYGHVEYTELLVKHCACVYNDLMQFTPLHEAAASKNR 223
DB 73 LLDHGAENVHCHDKDGTPLLAAYEGHVDVDDLLLEGADVDHNTDNGRTPLLAASMGH 132
QY 224 VEVCSLLISYADPTLLNCHNKSALDLAPTQPKERLAYEFKGSLLQ-AAAREADVTRIK 282
DB 133 ASVWNTLLFWGA-----AVD-----SIDSEGRVTSIASAOGNVVVR 170
QY 283 KHLSEMNVNFKIPQTHETALHCAASAPYKPKQICELLRRKANTNEKTEKFLPLHVAS 342
DB 171 TLLDRGLDENHRDDAGWTPLHMAA---PEGHRICEALIEOGARTNEIDNDGRIPFILAS 227
QY 343 EKAHNDVVEVYVYKHEAKYNALDNLGOTSLHRAAYCGHLQTCRLLLSYGCDPNISLOGFT 402
DB 228 QEGHYDCVOILLNENSDIDRGYDGRNALRVAALEGHRDIDVELLESHGAD-----277
QY 403 ALQMGNEVQQLLOEGISLGNSEADRLLEAAKAGDVETVKKLCVTQSVNCRDIEGROST 462
DB 278 -----VNCKDADGRPT- 288
QY 463 PLHFAAGYNRVSVVEYLLQHGADYHAKDGGVPLVPLHNACSYGHEVAELLVKHGAVNVA 522
DB 289 -LYILALENQLTMAEYFLENGANVEASDAEGRTALHVSQWGHMEMVQVLIAYHADVNA 347
QY 523 DLWKFTPLHEAAAKGYETCKLLLOHGADPTKKNRDGNTPDLVLVKGDTDIQDLLRGDAA 582
DB 348 DNEKRSALQSAWQGHVKKVQQLLIEHG-----A 375
QY 583 LLDAARKGGLARVKKLSPDNVNCRTDGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642
DB 376 VVD-----HTCNOGATALCIAAQEGHIDVQVLLLEGADPNHA 413
QY 643 DKGGLPLHNAASYGHVDVAALLIKYNACVNATDKWAFPTPLHEAAQKGRTQLCALLAHG 702
DB 414 DQFGRTAMRYAANKHNSQIKLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKY----467
QY 703 ADPTLK-NOGGOTPLDVSADVDVALLTAAMPSPALPCYKQVNLNVRSPGATADALSS 761
DB 468 QSLTIKSNSSGST-----GGGD-----MQPS-----LRGL--PNCPTHAFSS 502
QY 762 GPSPSSSL--SAASSLDNLGSGFSELSSVSSSGTEGASLEKKEVPGVDFSTIOFVRNL 819
DB 503 PSESPDSTVDQKSSLSNLSKSKSSSLRTTSSATAQT-----VPIDSFHNLSPTEQI 557
QY 820 GLEHLMDFPEREQI 833
DB 558 QOHSPLRSRSROSI 571
RESULT 9
LATA_LATMA STANDARD; PRT; 1401 AA.
AC P23631; Q25328;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-latrotoxin precursor.
OS Latrodectus mactans (Black widow spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Theridiidae; Latrodectus.
OX NCBI_TaxID:6924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tredcinguttatus; TISSUE=Venom gland;
```

```
RX MEDLINE-91031994; PubMed-1977615;
RA Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;
RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black
RW widow spider venom."
RL FEBS Lett. 270:127-131(1990).
RN [2]
RP PARTIAL SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE-91362695; PubMed-1888339;
RA Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;
RT "Structure of tryptic fragments of a neurotoxin from black widow
RW spider venom."
RL Bioorg. Khim. 17:437-441(1991).
CC -/- FUNCTION: Binds to high-affinity receptors that are localized in
CC the presynaptic plasma membrane of the nerve terminal. Binding of
CC toxin results in calcium influx into nerve terminals, activation
CC of synaptic vesicle exocytosis, and ATP depletion.
CC -/- PTM: ALPHA-LATROTOXIN MAY BE PROCESSED AT ITS C-TERMINUS.
CC -/- SIMILARITY: CONTAINS 21 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X55009; CAA38753.1; -.
DR PUR: S11527; S11527.
DR HSP: Q00420; IAWC.
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank; 20.
DR SMART: SM00248; ANK; 17.
DR PROSITE: PS50088; ANK_REPEAT; 11.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
KW Toxin; Neurotoxin; Repeat; ANK repeat; Venom; Signal.
FT SIGNAL 1 20
FT CHAIN 21 1401 ALPHA-LATROTOXIN.
FT REPEAT 490 521 ANK 1.
FT REPEAT 525 554 ANK 2.
FT REPEAT 559 589 ANK 3.
FT REPEAT 593 622 ANK 4.
FT REPEAT 626 656 ANK 5.
FT REPEAT 660 690 ANK 6.
FT REPEAT 695 723 ANK 7.
FT REPEAT 729 758 ANK 8.
FT REPEAT 762 791 ANK 9.
FT REPEAT 795 824 ANK 10.
FT REPEAT 828 857 ANK 11.
FT REPEAT 862 891 ANK 12.
FT REPEAT 895 924 ANK 13.
FT REPEAT 928 957 ANK 14.
FT REPEAT 971 1003 ANK 15.
FT REPEAT 1004 1033 ANK 16.
FT REPEAT 1035 1064 ANK 17.
FT REPEAT 1068 1097 ANK 18.
FT REPEAT 1101 1131 ANK 19.
FT REPEAT 1137 1166 ANK 20.
FT REPEAT 1170 1199 ANK 21.
SQ SEQUENCE 1401 AA; 156856 MW; B38A22083C142A98 CRC64;
Query Match 8.1%; Score 464.5; DB 1; Length 1401;
Best Local Similarity 23.2%; Pred. NO. 1.4e-20;
Matches 175; Conservative 94; Mismatches 274; Indels 211; Gaps 19;
QY 12 LQNGASVQARDGGLTPLHNACSFGEAEVNVLLLRHGADPNARD-NNVYTPLHEAAIKG 70
DB 513 LAKSTELNPDKKKGYTPIHVAADSGNAGIVNLLIRGVSVINSKTYHFLQTPHLAAQKGF 572
QY 71 IDVCIVLLQHGAEPTI-----RNTDGRATDLADPSAKAVLTGEYKKDELLESARSNGEEK 126
DB 573 VTFQRLME---SPEININERKDKGFTP-----LHYAIRGERI 608
```

```
QY 127 MMALLTPLNVNCHADGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAKDGLVPLHNA 185
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 609 LEAFLNQISIDVNAKSNLTGFHLAIKNDPVPASTLLGSKKVDINAVDENNTALHYA 668
QY 186 CSYGHVEVTE-----LLVKH-----GACVNAMD 209
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 669 AILGYLETTKQILNKELINANVSPGLLSALHYAILYKHDDVASFMRSSNVNKLAL 728
QY 210 WQFTPLHEAASNRVEVCSLLSYGADPTLLNCHNKSALDAPTQPKERLAYEFKGHSL 269
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 729 GGITPLHLAVIOGRKOILSLMPDIG-----753
QY 270 LOAREADVTRIKKLSLEVMVFKHPOTHE--TALHCAAAAPYPRKQICELLLRKGANI 327
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 754 -----VNIEQ-KDEKYPPLHLAAMSKYP---ELIQLLDQGSNF 789
QY 328 NEKTEFLTPLHVASEKAHNDVVEVVKHAKVNALDNLGOTSLHRAAYCGHLOTCRLL 387
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 790 EAKTNSGATPLHLATFKGSKQAALILLNNEVNRDTEGQMPIHGAAMTGLLDVAQAI 849
QY 388 SYGCDPNIIISLOGFTALQMGNEVQQLLEGISLGNSEADROLLEAAKAGDVETVKLC 447
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 850 S--IDATVV-----856
QY 448 VQSVNCRDIEGRQS-TPLHFAAGYNRVSVVEYLQHGADVHAKDGGVPLHNAACSYGH 506
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 857 -----DIEDKNSDTPNLAAQNSHIDVIKYFTDQGDADINTRNKGGLAPLAFSKKNL 909
QY 507 EVELLVKHGAVNVADLWKFTPLHEAAKGYEICKLQLLQHGADPTKKNRDGNTPDLV 566
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 910 DMVKYLPDKNAVYIADNDGNMFTYAVQNGHLNIVKYAMSEKOKFEWNSNTDN-----963
QY 567 KGDPTDIDQLLRGDAALDAAKKGLARVKKLSSPDNNVNCRDJTOGRHST--PLHLAAGYN 624
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 964 RRDCPNEECAISHFACVDAQVDRIETVKYFVG-----TLGNFAICGPLHQAARYG 1015
QY 625 NLEVAEYLLQH--GADVNAODKGLPLHNAASYGHVDVAALTIKYNACVNATDKWAF 681
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1016 HLDIVKYLVEEFLSVGSKTD---TPLCVASENGHFTVVOYLVSGAKVNHDCGNGMT 1071
QY 682 PLHEAAQGRGTOLCALLAHGADPTLKNQEGOTP 715
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1072 AIDRAITKHLQVQVFLAANGVDFRKNRSGTTP 1105

RESULT 10
V162_FOWPV
ID V162_FOWPV STANDARD; PRT; 603 AA.
AC Q9J569;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ankryrin-repeat protein FPV162.
GN FPV162.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zaak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- SIMILARITY: CONTAINS 14 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF198100; AAF44506.1; -.
DR HSSP; P25963; I1KN.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 12.
DR SMART; SM00248; ANK; 11.
DR PROSITE; PS50088; ANK_REPEAT; 11.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 23 53
FT REPEAT 57 87
FT REPEAT 91 120
FT REPEAT 124 155
FT REPEAT 159 189
FT REPEAT 193 223
FT REPEAT 227 257
FT REPEAT 261 291
FT REPEAT 295 325
FT REPEAT 329 362
FT REPEAT 366 397
FT REPEAT 401 428
FT REPEAT 432 467
FT REPEAT 504 533
SQ SEQUENCE 603 AA; 66542 MW; 8FEE3E464DB9D16D CRC64;

Query Match 7.3%; Score 420.5; DB 1; Length 603;
Best Local Similarity 25.5%; Pred. No. 2e-18;
Matches 143; Conservative 80; Mismatches 175; Indels 163; Gaps 17;

QY 145 KSTPLHLAAGY-NRVKIVQLLQHGADVHAKDGLVPLHNAAC-SYGHVEYTELLVKHGA 202
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 KDTPLHRAVLPDAVERIRMFVSGKADINVISDFKKTALHYAAKAKLATPEVLKTLIYLG 83
QY 203 CVNAMDLMQFTPLHEAASKNRVEVCSLLSYGADPTLLNCHNKSALDAPTQPKERLAY 262
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 NVNVTDMFESTPLHVAVQNGLEATKKLLDLGADP-----118
QY 263 EFKGSHLLQAAAREADVTRIKKLSLEVMVFKHPOTHEALHCAAAAPYPRKQICELLLR 322
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 -----NTRY-MNGOTPLHC-AAMVTPDGPVLRIIVE 148
QY 323 KGANINEKTEKFTPLHVASEKAH-NDVVEVVKHAKVNALDNLGOTSLHRAA-YCGLH 380
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 YGANVNALDNKHNTPLALAAELSNKTETIELGADVKKIKNDGITPLHLAAKSSDS 208
QY 381 QTCRLLSYGCDPNIIISLOGFTAL-----QMGNEVQQLLQEG--ISLGNSEADROLL 431
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 KTVETLILHGADVNAATCGSEGTPLHDAATSVELSN-TIEMLEIYGAEVNAASVGD---263
QY 432 EAAKAGDVETVKLCITVQSVNCRDIEGRQSTPLHFAA-GYNRVSVVEYLQHGADVHAKD 490
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 264 -----TPLHCAARSNPVHKLTKTLIAHGSNVNAVN 293
QY 491 KGGVPLHNAAC-SYGHVEVAELVKHGVNVADLWFTPLHFAAAAGKYEICK---LL 545
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 GISVTPHLATYSDNAATEALKVLIHGAEVNSVDIYQRTPMHYISRSYSOSLKTAVELL 353
QY 546 LQHGADPTKKNRDGNTPDLVQKGGDTDIQDLLRGDAALLDAAKKGLARVKKLSSPDNVN 605
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 VEHGADIEAKNVIGGTPLS-----SACNIE 379
QY 606 CRDTQGRHSTPLHLAAGYNLEVAEYLQHGADVNAODKGLPLHNAASYGHVDVAALL 665
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 380 -----YDLRIECLFIEYGADINTRIDETPLYSAIKY--PEIVNLL 419
QY 666 IKYNACVNATDKWAFPLHEA 686
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 MNYSASTNITNKSNTPLESA 440

RESULT 11
V222_FOWPV
```



RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RT Lehrach H., Reinhardt R., Yaspo M.-L.;  
 RL "The DNA sequence of human chromosome 21.";  
 CC Nature 405:311-319(2000).  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AB047783; BAB56136.1; -;  
 DR EMBL; AP001743; BAA95526.1; -;  
 DR HSP; P25963; IIRN.  
 DR GENE; HGNC:496; ANKRD3.  
 DR MIM; 605706; -;  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00023; ank; 10.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00248; ANK; 10.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50297; ANK\_REPEAT; 9.  
 DR PROSITE; PS50088; ANK\_REPEAT; 9.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 KW ANK repeat; Alternative splicing.  
 FT DOMAIN 22 286 PROTEIN KINASE.  
 FT REPEAT 485 514 ANK 1.  
 FT REPEAT 518 547 ANK 2.  
 FT REPEAT 551 580 ANK 3.  
 FT REPEAT 584 613 ANK 4.  
 FT REPEAT 617 647 ANK 5.  
 FT REPEAT 651 680 ANK 6.  
 FT REPEAT 684 713 ANK 7.  
 FT REPEAT 717 746 ANK 8.  
 FT REPEAT 750 780 ANK 9.  
 FT REPEAT 782 811 ANK 10.  
 FT NP\_BIND 28 36 ATP (BY SIMILARITY).  
 FT BINDING 51 51 ATP (BY SIMILARITY).  
 FT ACT\_SITE 143 143 BY SIMILARITY.  
 FT VARSPIC 278 325 MISSING (IN ISOFORM 2).  
 FT CONFLICT 714 714 M -> V (IN REF. 1).  
 SQ SEQUENCE 832 AA; 91610 MW; 5D8FFFD5F04F7ECB CRC64;  
 Query Match 6.8%; Score 390.5; DB 1; Length 832;  
 Best Local Similarity 28.1%; Pred. No. 2.2e-16;  
 Matches 141; Conservative 58; Mismatches 176; Indels 127; Gaps 13;  
 QY 228 SLLSYGADPTLLNCHNKSADIAPTPOLKERLAYEFKGLHSLQAAREADVTRIKKHLSL 287  
 Db 434 SLSLSPEREPEP-----STSDLGTTDQVKKK-----LVDAIVSGDTSKMLKIL-- 474  
 QY 288 EMVNFKHPQTHETALHCAAA-----SPYPRKQICELLRLKRGANINEKTEFLTPLHVA 343  
 Db 475 -----QPQVDVLDALDSGASLLHLAVEAQGECECAKWLILNNANPNLSNRGSGTPLHMAVE 528  
 QY 344 KAHNDVVVVVVKHKAVALDNLGQTSLLHRAAVCGHLOTCRLLSYGCDPNISLQGF- 402  
 Db 529 RRVGVVVELLLARKISVNAKDEQWTLHFAAQNQDSESTRLLLEKNASVNEVDFFGRTP 588

QY 403 ---ALQMGNN-VOQLQEGISLGNSEADQLLEAAKAGDVETVKKLCTVQSYNCRDIEG 458  
 Db 589 MHVACOHGOENIVRILLRRGDV-----SLOG 615  
 QY 459 RGS-TPLHFAAGYNRVSVVEYLL-QHGADVHAKDKGGLVPLHNSACSYGHVEAEILVKHG 516  
 Db 616 KQAWLPPLHYAAWQGHLPVVKLLAKAPGVSVNAQTLDGRTPLHLAAQGRHYRVARILIDL 675  
 QY 517 AVNVADLWKFPLHEAAAKGYEICKLILQHCADPTKKNRDGTPLDLVKDQDIDQL 576  
 Db 676 SDVNVCSLLAQPLHVAETGHTSTARLLHHRGAGKEAMTSDGYTALHL----- 724  
 QY 577 LRGDAAALLDAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHG 636  
 Db 725 -----AARNGHLATVK-----LLVEEK 741  
 QY 637 ADVNAQDKGLPLHNAASYGHVDVAALIKYNACVNAATDKWAFPLHEAAQKGRTOICA 696  
 Db 742 ADVLARGPLNOTALHLAAAHGSHVEELVSADV-IDLFEQGLSALHLAAQGRHAQTVE 800  
 QY 697 LLLAHGAD--PTLKNQEGQTP 715  
 Db 801 TLLRHGAHINLSLQKFGGHP 822  
 RESULT 13  
 V244\_FOWPV STANDARD; PRT; 568 AA.  
 AC Q9J4Z6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Putative ankryrin-repeat protein FVP244.  
 GN FVP244.  
 OS Fowlpox virus (FPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Avipoxvirus.  
 OX NCBI\_TaxID=10261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20193820; PubMed=10729156;  
 RA Afonso C.L., Tullman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
 RT "The genome of fowlpox virus.";  
 RL J. Virol. 74:3815-3831(2000).  
 CC -!- SIMILARITY: CONTAINS 12 ANK REPEATS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF198100; AAF44588.1; -;  
 DR HSP; P42773; IIRH.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam; PF00023; ank; 10.  
 DR SMART; SM00248; ANK; 6.  
 DR PROSITE; PS50088; ANK\_REPEAT; 7.  
 DR PROSITE; PS50297; ANK\_REPEAT; 1.  
 KW Hypothetical protein; Repeat; ANK repeat.  
 FT REPEAT 40 69 ANK 1.  
 FT REPEAT 144 173 ANK 2.  
 FT REPEAT 177 206 ANK 3.  
 FT REPEAT 210 239 ANK 4.  
 FT REPEAT 272 302 ANK 5.  
 FT REPEAT 306 336 ANK 6.  
 FT REPEAT 340 370 ANK 7.  
 FT REPEAT 374 403 ANK 8.  
 FT REPEAT 407 437 ANK 9.  
 FT REPEAT 441 471 ANK 10.



```
FT REPEAT 473 502 ANK 11.
FT REPEAT 571 602 ANK 12.
SQ SEQUENCE 668 AA; 76200 MW; F3C5BA076023791B CRC64;

Query Match
Best Local Similarity 27.3%; Score 380.5; DB 1; Length 668;
Matches 153; Conservative 95; Mismatches 192; Indels 121; Gaps 26;

QY 186 CSYGHVEYVELLV-----KHGA--CVNAMDLDWQFTPLHAEAAKRNVEVCSSLISYGDPT 238
DB 11 CRGNYIELSAITNNYLNHKGACNOCEN--ESIPETAIHQALQRLDIDIVKELIO--QNPX 66
QY 239 LL-----NCHNKSADLATPPLKARLAEYFKGHSLLQAAAREADVTRIKKHLSELMVFNFKH 294
DB 67 LIYVTDHRNSTLHICITPNVMD-----IVISLTVDCDIILDIKYASI--ILN-KH 115
QY 295 POTHETALHCAAAASPYRKKQICELELLLRKANINE-----KTKFELTPLHVAASEKAHND-- 348
DB 116 -----KLGECACIHLKREGISGNEISYNKINKSTEYMKLIKERTQODEL 158
QY 349 -VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIIISLOGFTALQ-- 405
DB 159 LIAEMLLKKGIDVNAKDVYCRTPHIAAERGNTKMNVLNLLSYGADVNIITLDDLSVLEYA 218
QY 406 MGENVQOQLLQEGISLGN--SEADROLLEAAKAGDVETVKKLC-TVQSVNCRDIEGRQSTP 463
DB 219 VDSKNIDTITKALIDNRSNKNKNDLSLKAIRNTDLETSLLYDSGFSVNSIDV--YKNTP 276
QY 464 LHFAAGYNRVS-VWEYLLQHGADVHAKDGLVPLHNAACSYGH-VYEAELLVKGAVNVV 521
DB 277 LHVAVQAPSLRVLKPLRGIDVNAKNIKGETPLYLAKNGYDPTNIRTLIMRGADVNA 336
QY 522 ADLWFTPLHFAAAAGKYEICKL-LLOHGADPTKKNRGDPTPLDLVKGDDTDIDQLLRGD 580
DB 337 ADSLYITPLHQASTLDRYKDTVITLLEGA----- 366
QY 581 AALLDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 640
DB 367 -----NVNARDYCK--TPHIAAVRNNVVINTLLDYGDADIE 402
QY 641 A-ODKGGLIPLHNAASYG---HVDVAALLIKYNACVNTDKWAFPLHFAAOKG-RTOLC 695
DB 403 ALSQKIGTV-LH-FALYGNFPMYSVKT-LIDRGANVNSKNKYLSTPLHYACKNKCKPEVI 459
QY 696 ALLLHAGADPTLKNQBGOTPL 716
DB 460 KMLLDNGADVNAINIRNQPL 480

RESULT 14
V024_F0WVPV STANDARD; PRT; 596 AA.
AC Q9J5H7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative ankryin-repeat protein FPV024.
GN FPV024.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus."
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: CONTAINS 13 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF198100; AAF44368.1; -.
CC HSSP; P42773; 1IHB.
CC InterPro; IPR002110; ANK.
CC Pfam; PF00023; ank; 12.
CC SMART; SM00248; ANK; 9.
CC PROSITE; PS00088; ANK_REPEAT; 7.
CC PROSITE; PS0297; ANK_REPEAT; 1.
CC Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 5 34 ANK 1.
FT REPEAT 37 66 ANK 2.
FT REPEAT 68 96 ANK 3.
FT REPEAT 99 128 ANK 4.
FT REPEAT 130 158 ANK 5.
FT REPEAT 162 191 ANK 6.
FT REPEAT 195 225 ANK 7.
FT REPEAT 229 258 ANK 8.
FT REPEAT 262 291 ANK 9.
FT REPEAT 295 324 ANK 10.
FT REPEAT 326 355 ANK 11.
FT REPEAT 359 389 ANK 12.
FT REPEAT 394 423 ANK 13.
SQ SEQUENCE 596 AA; 67938 MW; F813C43228907D07 CRC64;

Query Match
Best Local Similarity 25.2%; Score 368; DB 1; Length 596;
Matches 135; Conservative 83; Mismatches 177; Indels 140; Gaps 19;

QY 146 STPLHLAAGYNRVKIVOLLQHGADVHAKDGLVPLHNAACSYGHVEYVELLVKHGACVN 205
DB 39 STPLVAVNSDDIDIVKFLDNGADI---NKCSPPHLKALNGLNVEWKLVDHGADIE 95
QY 206 AMDLWQFTPLHFAAASKRVEYCSLLSYGADPTLL-----NCHNKSADLATPPLQK 259
DB 96 KVLGN-SPLYLALCKRNTITKYLLERGADPTLFTNYCDIYKIPIDI----- 145
QY 260 LAYEFKHSLLQAAAREADVTRIKKHLSELMVFNFKHPOTHETALHCAAAASPYRKKQICE 319
DB 146 ----FK-----ILIKYKVSLSNIQNSHFKTPYY--AIKC---TNP-----LIK 181
QY 320 LLRKGANINEXTKFELTPLHVAASEKAH--DVVEVVVKHEAKVNALDNLGQTSLHRAAY 376
DB 182 LLENNASLT--IPEGNNHYLITAVKHNCDISILRLIKYGVVPVNEQDLETSLSHYCVS 239
QY 377 CGHLQTCRLLSYGCDPNIIISLOGFTALQMGNEVQOQLLQEGISLGNSEADROLLEAAKA 436
DB 240 AGKHDILKLLDYDADPNI-----TDSCLGTP-----LHYAVSRNDIATITLLIE--KG 286
QY 437 GDV----ETVKKLCITVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 492
DB 287 ANVINHNQDI-----DTVLTAVGNRNKILINLLMGANTPLKSRN 328
QY 493 GLVPLHNAACSYGHVEAELLVKGAVNVADLWKFTPLHFAAAK--GKYEICKLLOHGAD 551
DB 329 PLI--HKAETKDINILSEILNHGAENVINREGVTPLYIAITFMQIKFAKLLRYG 386
QY 552 PTKKNRQDNTPLDLVKGDDTDIDQLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQ 611
DB 387 PNMKN-----ES 393
QY 612 RHSTPLHLAAGYNLEVAEYLLQHGADVNAODKGLIPLHNAASVGHVDVAALLI 666
DB 394 NENTPLHGAILSNRLDSVELLSYNSVNDVHSINKLGHPTPL-SCINVISDKIATIII 447

RESULT 15
DAPK_HUMAN
ID DAPK_HUMAN STANDARD; PRT; 1431 AA.
AC P53355;
```



Qy 875 -----NPYLTNTSGTILIDLSPDDKEFQSVEEEMQSTV 910  
Db 917 NRFGNDLHISNKLFLVDAGSG-----SNDMKVLRNHLQEIERSQI 956

Search completed: February 12, 2003, 03:33:01  
Job time : 18.6325 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 12, 2003, 03:26:09 ; Search time 41.3675 Seconds  
(without alignments)  
5478.980 Million cell updates/sec

Title: US-09-843-159B-3

Perfect score: 5769

Sequence: 1 GFGKDVVEYLLONGASVQA.....AYPEYLITYQIMRPEGWVDG 1100

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mmc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	3335.5	66.5	1181	5	Q9XZ37
2	3335.5	66.5	1181	5	Q9VBP3
3	875.5	15.2	1726	11	Q8VC68
4	875.5	15.2	4377	4	Q12955
5	873.5	15.1	1943	11	Q61307
6	867.5	15.0	1136	6	Q9N180
7	864.5	15.0	2622	11	O70511
8	863.5	15.0	1762	11	O88521
9	842	14.6	1719	4	Q13768
10	841.5	14.6	843	11	P97582
11	840.5	14.6	1856	4	Q99407
12	839.5	14.6	1848	11	Q61302
13	836	14.5	1549	5	Q9V4B1
14	834	14.5	1549	5	Q24241
15	820	14.2	1159	5	Q9NCP8
16	814.5	14.1	1809	5	Q17487

17	814.5	14.1	1867	5	Q17486
18	814.5	14.1	2039	5	Q17489
19	810.5	14.0	1815	5	Q17488
20	810.5	14.0	6994	5	Q17343
21	807.5	14.0	1786	5	Q17344
22	799	13.8	2443	5	Q9VSA2
23	793	13.7	1009	5	O8SWY2
24	629.5	10.9	1088	4	Q13484
25	628	10.9	791	4	Q9UFA4
26	601	10.4	1913	5	Q9GRV5
27	584.5	10.1	1619	5	O9NGR8
28	584	10.1	1453	5	Q9VMR4
29	554.5	9.6	1098	11	Q61304
30	545.5	9.5	1411	5	Q02989
31	534.5	9.3	1599	11	Q99NH0
32	527	9.1	743	4	Q8TB46
33	520	9.0	1395	5	Q9XZC0
34	519	9.0	1282	5	O8TF48
35	519	9.0	2119	5	O9VAU5
36	519	9.0	2119	5	O8TG99
37	514	8.9	934	16	O83807
38	509	8.8	1021	13	Q8UVC1
39	500.5	8.7	1486	4	Q8TEF1
40	499.5	8.7	833	4	Q96I86
41	496.5	8.6	1062	11	O89019
42	491.5	8.5	1062	11	O88849
43	489.5	8.5	1188	4	Q9H288
44	486	8.4	895	4	Q9Y488
45	486	8.4	1065	4	Q9Y283

## ALIGNMENTS

### RESULT 1

Q9XZ37 ID Q9XZ37 PRELIMINARY; PRT: 1181 AA.  
AC Q9XZ37;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 127.9 kDa protein.  
GN BCDA:LD22548 OR CG4719 OR CGI7487.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,  
RA Abayani A., Arcaina T.J., Baxter E., Blazej R.G., Butenhoff C.,  
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,  
RA Gallego R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,  
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,  
RA Park S., Sequeira A., Sethi H., Snir E., Svirska R.R., Weinburg T.,  
RA Celniker S.E.  
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AF132196; AAD34784.1; -.  
DR HSSP; Q00420; IAWC.  
DR FlyBase; FBgn0027508; BcDNA:LD22548.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR001660; SAM.  
DR Pfam; PF00023; ank; 17.  
DR Pfam; PF00536; SAM; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 15.  
DR SMART; SM00454; SAM; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 14.  
DR PROSITE; PS50297; ANK\_REPEAT; Repeat.  
KW ANK repeat; Hypothetical protein; Repeat.  
SQ SEQUENCE 1181 AA; 127884 MW; 4E2BF31549BBA2D9 CRC64;

Q17486	caenorhabdi
Q17489	caenorhabdi
Q17488	caenorhabdi
Q17343	caenorhabdi
Q17344	caenorhabdi
O9VSA2	drosophila
O8SWY2	drosophila
Q13484	homo sapien
Q9UFA4	homo sapien
Q9GRV5	caenorhabdi
Q9NGR8	drosophila
Q9VMR4	drosophila
Q61304	mus musculu
Q02989	latrodectus
Q99NH0	mus musculu
Q8TB46	homo sapien
Q9XZC0	latrodectus
O8TF48	drosophila
O9VAU5	drosophila
O8TG99	drosophila
O83807	treponema p
Q8UVC1	brachydanio
Q8TEF1	homo sapien
Q96I86	homo sapien
O89019	mus musculu
O88849	mus musculu
Q9H288	homo sapien
Q9Y488	homo sapien
Q9Y283	homo sapien

Query Match		66.5%; Score 3835.5; DB 5; Length 1181;
Best Local Similarity		65.3%; Pred. No. 7 4e-237;
Matches 740; Conservative 156; Mismatches 182; Indels 55; Gaps 12;		
QY	1	GFGKDVYELLQNGASVQARDGGGLIPLHNACSFGEAEVWVLLRHGADPNARNWNT 60
DB	66	GYGRREVVEFLNNGASIQACDEGLHPLHNCSCFGEAEVWVLLKAGASPTTNDNNT 125
QY	61	PLHEAAIKKIDCVILVLOHGAETIRNIDGRTALDADPSAKAVLTGEYKDELLESAR 120
DB	126	PLHEAAKGVKDVCLALQGANHTIRNSEQTPLELADENRPTVLYRDELEAAR 185
QY	121	SGNEEKWALLPLNVNCHASDGRKSTPLHAAAGYNRKIVOLLHGHGADVHAKDKGLV 180
DB	186	SGAEDRLIALPLNVNCHASDGRSTPLHAAAGYNRIGIVELLANGADVHAKDKGLV 245
QY	181	PLHNACSGHYEYVELLVKHGACVNMADLWQFTPLHEAAKSNRVEVCSLLLSYGADPTLL 240
DB	246	PLHNACSGHYFDVTKLIQAGANVNANDLWAFPLHEAAKSNRVEVCSLLLSRGADPTLL 305
QY	241	NCHNKAIDLAPTPOLKERLAYEFKCHSLLOAAREADVTIRKKHLSLEVMNFKHPOTHET 300
DB	306	NCHSKSAIDAAPTRERIRAFKYGCHLLDADCRKCDVSRANKLVCAEIVNFVHPYTGDT 365
QY	301	ALHCAASPYPRKQICELLRRKANINKEKTEFLPLHVAESEKAHNDVVVYVVRKEAV 360
DB	366	PLHVAVSPDKRQKLMELLTRKGLSLLNEKNAFLPLHLAELLHYDAMEVLLKQAKV 425
QY	361	NALDNLGQTSLHRAAYCGHLOTCRLLSLCYGCDNIIISUGFTALONGNENVOQLLEGIS 420
DB	426	NALDSLGQTPHRCARDEQAVRLLLSYAADPNIVSLEGLTAAQASDVLKLLK--- 479
QY	421	LGNSEADROLLEAAKAGDVETVKKLCTVQ--SYNCRDIEGROSTPLHFAAGYNRVSWVEY 478
DB	480	-NPPDSETHLEAAKAGDLDTVRIVLNNPISVNCRDLDGRHSTPLHFAAGYFRVPPVQF 538
QY	479	LLOHGDADVHAKDKGGLVPLHNACSGHYEYAEVLLVKGAVNVNADLWKTPLHEAAAKG 538
DB	539	LLEHGAEVYAADKGLVPLHNACSGHYEYVELLVKHGANVNVSDLWKTPLHEAAAKG 598
QY	539	YEICKLLLOHGDPTKKNRDNTPDLVKGDDPDIDQLLRGDALDLDAKKGCLARVKKL 598
DB	599	YDICKLLKHGADPMKKNRDNATPDLVKSDDHVDLELRGSPALLDAKKGKLARVQL 658
QY	599	SSPDVNCRTQGRHSTPLHAAAGYNNLEVAEYLLHGHGADVNAQDKGGLIPLHNAASYGH 658
DB	659	VTEPINCRAOGRNSTPLHAAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNAASYGH 718
QY	659	VDVAALLIKYNACVNTDKWAFPLHEAAKGRGTQICALLAHGADPTLKNQEGQTPDL 718
DB	719	LDIAALLIKHTVYVNTDKWGTPLHEAAKGRGTQICALLAHGADAYMKNQEGQTPIEL 778
QY	719	VSADDYSALLTAAMPSS---ALPSCYKPOVLNGVRSPGATADALSSGSPSSSS--LSAA 772
DB	779	ATADDVKLLQDAMATLSQOAL-SASTQSLTSSSPADPATAAAACPTSSSSSAILSP 837
QY	773	SSLDNLGSGSEISSV---VSSS-----GTEGASSLEKKEYPGVYD--FSIT 813
DB	838	TETVLLPTGASMLISVPVPLPLSSSTRISPAQAEANGAEGSSDDL-LPDADTITNVS 895
QY	814	QFVRNLGLEHMDIFEREQITLDVVMGMHKEKELKEIGINAYGRHKLIKVERLISGOQG 873
DB	896	GFLSSQOHLHLELFEREQITLDLAEHGHDDLLKQGVSAIGFRHKLKIAQLRS--- 951
QY	874	LNPLYLTNTSGG-----TILIDLSPDDKFOSEFMOSTVREHRDGGHAGIENRYN 927
DB	952	-----TTGIGNNVNLCITLLVDLLPDDKEFVAVBEEMQATTIREHRDNGAQGYFTRYN 1003
QY	928	ILKIQYCNKKLWERYTHRRKEYSEENHNHANRMLFHGSPFFVNAIITHKGFDERHAYIG 987
DB	1004	IIRVQYVONKLVYAHRRQETAEENFLQSNRMLFHGSPFFNAIVORGFDERHAYIG 1063
QY	988	MFGAGIYFAENSSKSNQYVYGGTCCPVHKDRSCYICHRQLLFCRYTLGKSLQESAM 1047

DB	1064	MFCAGIYFAEHSSKSNQYVYGGTCCPVHKDRSCYICHRQLLFCRYTLGKSLQESAM 1123
QY	1048	KMAHPPGHHSVTGRPSVNGALAEAYIVYRGQAYPEYLITYQIMRPEGWVG 1100
DB	1124	KMAHPPGHHSVYGRPSAGGLHFAEYVYRGQAYPEYLITYQIVKDDSSSG 1176
RESULT 2		
Q9VBP3		
ID	Q9VBP3	PRELIMINARY; PRT: 1181 AA.
AC		
DT	01-MAY-2000 (TREMBLrel. 13, Created)	
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	BCDNA:LD22548 protein.	
GN	BCDNA:LD22548 OR CG4719 OR CG17487.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BERKELEY;	
RX	MEDLINE=20196006; PubMed=10731132;	
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,	
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,	
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Bottier P.,	
RA	Burkitt K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA	Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,	
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,	
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,	
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,	
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	
RA	Shue B.C., Sidenkian I., Simpson M., Skupski M.P., Smith T.,	
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,	
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,	
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;	
RT	*The genome sequence of Drosophila melanogaster.*;	
RL	Science 287:2185-2195(2000).	
DR	EMBL; AF003752; AAF56487.1;	
DR	HSSP; Q00420; 1AWC.	
DR	FlyBase; FBgn0027508; BCDNA:LD22548.	
DR	InterPro; IPR002110; ANK.	
DR	InterPro; IPR001660; SAM.	
DR	Pfam; PF00023; ank; 17.	
DR	Pfam; PF00536; SAM; 1.	
DR	PRINTS; PR01415; ANKYRIN.	
DR	SMART; SM00248; ANK; 15.	

[illegible]

```
Db 340 NVPDVTNDYLTALHVAACHGCHKYK---VAKVLLDKKASPNKALNGFTPLHIAACKNNRI 396
Qy 348 DVVEVVVHKAQVNDNLGOTSLHRAAYCGHLCQTRLLSYGCDPNITISLGFTALQMG 407
Db 397 RWELLKHGASIQAVTESGLTPIHVAAPMGHVNIVSQLMHGASPNITNVRGETALHMA 456
Qy 408 -----NENVOOLLOEGISIGNSEADRO---LLEAAKAGDVETVKKLCV-OSVNCRDIEGR 459
Db 457 ARSGQAEVRYLVQDGAQVEAKAKDDQTPHISARLGRADIVQOLLOQASPNATTS- 515
Qy 460 QSTPLHFAAGYNRVSVVEYLQHGADVIAKDKGLVPLHNAACSYGHEVEAEVLLKHGAV 519
Db 516 -YTPHLAAREGHEDVAFAFLDHGASLSITTKKGTPLHVAAYKGLVAVSLLOKSASP 574
Qy 520 NVADLWKTPLHEAAKGYEICKLLQHGADPTKKNRGDNTPLDL-VKDGTDIDQDLL- 577
Db 575 DAAGSGLTPLHVAHYDNQKVALLELDQGASPHAAKNGYTPHIAAKKNQMDIATSL 634
Qy 578 -----RGDAALLDAKKGCLARYKLLSSPD-NVNCRDTOGRHSTPLHLAAGYNN 625
Db 635 EYGADANAVTROGIAVSLHAAQEGHVDVMVLSLLSRNANVLSNKS- -LTPLHLAAQEDR 692
Qy 626 LEVAEYLLQHGADVNAQDKGLIPLHNAASGHVDVAALLIKYNACVNAATKWAFTPLHE 685
Db 693 VNAEVLVNOGAHVDAQTKMGYTPPHVGHYGNIKIVNLFLLQHSKAVNAKTKNGYALHQ 752
Qy 686 AAKGRTOLCALLHAHADPTLKNQEGOTPLDLVS-----ADVSALLTAAMPSPALP 738
Db 753 AAOQGHTHIIVLNONNAPSNELTVNGNTALAIARRLGIVSVVDITKVVTEIMTTIT 812
Qy 739 SCYK---POVLNGV--RSPGATADALSS-----GP-----SSPSSLSAASSLDNL 778
Db 813 EKHKNNVPETMNEVLMDSDDEGDAITGTDKYLGPQDLKELGDDSLPAEGYVGSFGAR 872
Qy 779 SCSFSELSVVS-----SSGTGASLSKEKVPVGVDFSTQTVRNGLGHLMDIFEREQI 833
Db 873 SASLSFSDSRSTYLRSSYARDSMWIBELLYPSKEQHLT-FTREFDSDSLR-YSWAAD 930
Qy 834 TLVDL-----VEMGHKELKEIGINAYG-----HRHKLI-----KGVRLIS 869
Db 931 TLDNVLNVSSPVHSGF--LVSPMVDARGSMGRSHHGMRIIIPRKCTAPTRICRLVK 988
Qy 870 GOOGLNPLYLTNTSGSITLIDLSPDKEFQS-VVEEMQSTVREHRDGGHAGGIFNRNYI 928
Db 989 RHKLANPPPMVEGEGLASRLVEMGPAGAQLGPVIVEIP-----HFGSMRGKERE 1038
Qy 929 LKIOKVCNKKLWE--RYTHRRKEVE-----ENHNHANRMLFHGSPFNALIIKHGEDER 981
Db 1039 LIVLSENGETWKEQFDSKNEDELLNGMDEELDSPEEL---GTRICRIITKDFPQ- 1094
Qy 982 HAYIGCMFGAGIYFAENS---SKSQYVYVIGGG 1012
Db 1095 -----YFAVVSRIKQESNQ--IGPBG 1114
RESULT 4
ID Q12955
AC Q12955 PRELIMINARY; PRT: 4377 AA.
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ankyrin G.
GN ANK-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN STEM;
RX MEDLINE=95138209; PubMed=7836469;
```

```
RA Kordeli E., Lambert S., Bennett V.;
RT "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RL at the axonal initial segment and node of Ranvier.";
RN J. Biol. Chem. 270:2352-2359(1995).
RP [2]
RC SEQUENCE FROM N.A.
RA TISSUE=BRAIN STEM;
RC Carpenter S.S.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U13616; AA64834.1; -.
DR HSSP; P55273; IBI8.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS00088; ANK_REPEAT; 21.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;
```

```
Query Match 15.2%; Score 875.5; DB 4; Length 4377;
Best Local Similarity 30.5%; Pred.No. 2.1e-46;
Matches 273; Conservative 126; Mismatches 362; Indels 135; Gaps 21;
Qy 3 GRKDVYVLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 62
Db 85 GHVEVVSLLQREANVDAATKKGNTALHIALASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 144
Qy 63 HEAAIKGIDVCIIVLQHGAEPTIRNTDGRYALDLA-----DPSKAVUTGEYK---KDE 114
Db 145 YMAQENHLEVKVFLDNGASQSLATEDGFTPLAVALQGHQDVVSVLLLENDTKGKVRLP 204
Qy 115 LLESARSGNEEKMMALPLNVNCHASDCKSTPLHLAAGYNRKIVTOLLQHGADVHAK 174
Db 205 ALHIAARKDDTKAAALLQNDNNDVADVESKSGFTPLHIAAHYGNINIVATLLNRAAAVDFT 264
Qy 175 DKGDVLPLHNACSYGHEVTELLVKKGACVNAQMLQVFTPLHEAAKSNRVEVCSLLSYG 234
Db 265 ARNDITPLHVASKRGNNANVMKLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVMELLDRA 324
Qy 235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKCHSLLOAREADVTRIKKHLSELMVNFH 294
Db 325 A-PILSKTKN-----GLSLPHMATQGD-----HLNCVQLLOH 356
Qy 295 -----PQTHETALHCAASPYPKRKOICELLRLKGANINEKTEFTPLHVAEKAHN 347
Db 357 NVPDVTNDYLTALHVAACHGCHKYK---VAKVLLDKKASPNKALNGFTPLHIAACKNNRI 413
Qy 348 DVVEVVVHKAQVNDNLGOTSLHRAAYCGHLCQTRLLSYGCDPNITISLGFTALQMG 407
Db 414 KVMELLKHGASIQAVTESGLTPIHVAAPMGHVNIVSQLMHGASPNITNVRGETALHMA 473
Qy 408 -----NENVOOLLOEGISIGNSEADRO---LLEAAKAGDVETVKKLCV-OSVNCRDIEGR 459
Db 474 ARSGQAEVRYLVQDGAQVEAKAKDDQTPHISARLGRADIVQOLLOQASPNATTS- 532
Qy 460 QSTPLHFAAGYNRVSVVEYLQHGADVIAKDKGLVPLHNAACSYGHEVEAEVLLKHGAV 519
Db 533 -YTPHLAAREGHEDVAFAFLDHGASLSITTKKGTPLHVAAYKGLVAVSLLOKSASP 591
Qy 520 NVADLWKTPLHEAAKGYEICKLLQHGADPTKKNRGDNTPLDL-VKDGTDIDQDLL- 577
Db 592 DAAGSGLTPLHVAHYDNQKVALLELDQGASPHAAKNGYTPHIAAKKNQMDIATSL 651
Qy 578 -----RGDAALLDAKKGCLARYKLLSSPD-NVNCRDTOGRHSTPLHLAAGYNN 625
```





QY 115 LLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAK 174  
DB 188 ALHIAARKDDTRAAALLLQNDNADVESKSGFTPLHIAAHYGNINVATLLNRAARAVDT 247  
QY 175 DKGDLVPLINACSYGHEVTELLVHKGACVNDLWQFTPLHEAASKNRVEVCSLLSYG 234  
DB 248 ARNDITPLHVASKRGNNANMKLLDGRGAKIDAKTRDGLTPLHCGARGSEHQVVMELDRS 307  
QY 235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKHLSLEWVNFKH 294  
DB 308 A-PILSKTKN-----GLSPHUMATQGD-----HLNCVQLLQHG 339  
QY 295 -----POTHETALHCAASPPYKQICELLKRGKGANINEKTEFLPLHVAASEKAHN 347  
DB 340 NYPVDVDDTYLTLHVAACHGKHYK---VAKVLLDKKASPNKALNGFTPLHIAACKNRI 396  
QY 348 DVEVVVHKEAKVNDLNGQTSRLHRAAYCGHLOTRCLLSYGCDDNITISLQGTALOMG 407  
DB 397 RVWELLKHGASITQAVTESGLTPIHVAAPMGHVNIVSQLMHGASPTTNRGETALHMA 456  
QY 408 -----NENVOOLLOGBISLGNSEADRO---LLEAAKAGDVETVKKLCIV-OSVNCRDIEGR 459  
DB 457 ARSGOAEVRYLVQDGAOVEAKAKDDQTPHISARLGKADIVQOLLQOGASNAATTS- 515  
QY 460 QSTPLHFAAGYNRVSYVEYLQHGADVHAKDKGGLVPLHACSYGHEVVAELLVKGAVV 519  
DB 516 -YTPLHIAAREGHEDYAAFLDHDGASLSITTKGFTPLHVAAYKGLVASLLLOKSASP 574  
QY 520 NVADLWKFPLHEAAKGYEICKLLOHGADPTKKNRDNTPDL-VKGDGTDIDOLL- 577  
DB 575 DAAGKSGLTPLHVAAYHYNQKVALLDQGASPHAAKNGYTPHIAAKKNQMDIATSL 634  
QY 578 -----BGDALLDAAKGGLARVYKLLSSPD-VNVCRTDQGRHSTPLHLAGYNN 625  
DB 635 EYGADANVTRGIAVSHLAAQEGHVDVMSLLSRNANVNLNKS- -LTPHLHAAQEDR 692  
QY 626 LEVAEYLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHE 685  
DB 693 VNVAEVLVNOGAHVDAQTGMVGTPLHVGHYGNIKIVNELLQHSKAVNAKTNGYTLHQ 752  
QY 686 AAKGRTQLCALLAHGADPTLKNQEGTPLDLS-----ADVSALLTAMPSPALP 738  
DB 753 AAOQGHITHIINVLQNNASPNELTVNGNTALAIARLGYISVVDTLKVYTEIMTTIT 812  
QY 739 SCYK---POVLNGVRSPGATADALSGSPSSLSAASSLDNLS-GSFSELSSVSSSGT 794  
DB 813 EKHKNVPTMNEVLD-----MSDEVKASAPEKLSGEY-----ISDGE 853  
QY 795 EGASSLEKKEVPQVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKEIG 850  
DB 854 EG-----EDAITGDTKYLGPQDLKELG 876

RESULT 6  
Q9N180  
ID Q9N180 PRELIMINARY; PRT: 1136 AA.  
AC Q9N180;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Ankyrin 1 (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRED JAPANESE BLACK; TISSUE=BONE MARROW;  
RA Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;  
RT "Cloning of bovine erythrocyte ankyrin."  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF222766; AAF61702.1; -.  
DR HSSP; Q00420; IAWC.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00023; ank; 24.  
DR Pfam; PF00791; ZU5; 1.  
DR SMART; SM00248; ANK; 22.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS00088; ANK\_REPEAT; 20.  
DR PROSITE; PS0297; ANK\_REPEAT\_REGION; 1.  
KW ANK repeat; Repeat.  
FT NON\_TER 1136 AA; 1136  
SQ SEQUENCE 1136 AA; 12182 MW; 07CD8B1568ACE8C7 CRC64;  
  
Query Match 15.0%; Score 867.5; DB 6; Length 1136;  
Best Local Similarity 26.6%; Pred. No. 8.7e-47;  
Matches 283; Conservative 157; Mismatches 375; Indels 249; Gaps 28;  
  
QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSFGEHAEVYVNLIRHGADPNARDNNMYTPL 62  
DB 88 GODEVVRVLYNRYGANNVNAQSQKGFPLYMAQENHLEVVKFLLENGANQNVATEDGFTPL 147  
QY 63 HEAAIKGKTDCVILVLOHGAETPI-----RNTDGTALDLADPSAKAVUTGEYKDK 113  
DB 148 AVALOQGHENVVAHLINYGTKGVRLPALHIAARNDDTRTA-----VLLQNDPNP 198  
QY 114 ELLESARSNGNEEKMMALLTPLNVNCHASDGRK-----TPHLHLAGY 155  
DB 199 DVL--SKTG-----FTPLHIAAHYENLVNAQLLNNGASVNFPPQNGITPLHIASRR 248  
QY 156 NRVKIVQLLQHGADVHAKDKGDLVPLHACSYGHEVTELLVHKGACVNDLWQFTPL 215  
DB 249 GNVIVRLLLDGRGAQIETRTKDELTPHCAARNHRLRISEILLDHGAPQAKTKNGLSPI 308  
QY 216 HFAASKNRVEVCSLLSYG---DPTL-----LNC-HNKA---IDLAPTOLKER 259  
DB 309 HMAAGDHLDCVRLLYQNAEIDITDILTPHVAAGHHRVAKVLLDKGAKPNRS-- 366  
QY 260 LAVEFKGHSLLQAAAREADVTRI-----KKHLSLEWVNEFKHPOTHETALHCAASPYPKR 314  
DB 367 ---ALNGFTPLHIAACKNHIRVWELLKMGASIDAVT-----ESGLTPLHVASPMGHP--- 416  
QY 315 QICELLKRGANINEKTEFLTPHVAASKAHNDVVEVYVVKHEAKVNDLNGQTSRLHRA 374  
DB 417 PIVKSLLOREASPNVSNVKTETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPHCA 476  
QY 375 AYCGHLOTRCLLSYGCDDNITISLQGTALOMGNENVOOLLQGBISLGNSEADRLLEAA 434  
DB 477 ARIGHTNMVYKLLLENANPNLATAGTTPHIA-----AA 510  
QY 435 KAGDVETVKKLCTVO-SVNCRDIEGRQSTPLHFAAGYNRVSYVEYLQHGADVHAKDKG 493  
DB 511 REGHVETALALLEKEASQTCMTKKG--FTPLHVAAYKGVKMAELLEDHDAHNAAGKSG 568  
QY 494 LVPLHNACSYGHEVVAELLVHKGAVVNVADLWQFTPLHEAAAKGYEICKLLOHGAADPT 553  
DB 569 LTPHVAHVHNLVDVRLLPGRGSPHSPALNGYTPHIAAKQNLQVARSLLQYGSAN 628  
QY 554 KKNRGNTPLDL-VKGDGTDIDOLLRGDAALLDAKKGCLARVKKLSSPDNVNCRDQGR 612  
DB 629 AESVOGVTPLHIAAQEGHAEWVALL-----LSKQANGNLGNKSG- 667  
QY 613 HETPLHLAGYNNLEVAEYLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACV 672  
DB 668 -LTPHLHVAQEGHIPVADVLKIHGTVTDATRMGYTPHVAHYGNIKLVKFLQLQHKADV 726  
QY 673 NATDKWAFPLHEAAKQGTQLCALLAHGADPTLKNQEGTPLDL-----VSADDSVA 726  
DB 727 NAKTKLGYSPLHQAQOQGHDTIVILLKHGASPNVSSNGTTPLAIAKRLGYISVTDVLK 786  
QY 727 LITAAMPSPALPSCYK---POVLN-----GYRSPGATA----- 756  
DB 787 VVTDETTVOLITDKHRMSPFETVDEILDVSEDEGEDLGPRAERPDARDVEEKELDFVP 846

```

QY 757 --DALSSGSPSP-----SSLSAASLDNLGSGFSELSSVSSSGTEGASSLEKKEVP 806
DB 847 KLDQVESPALPRICVPTETVIRADPDQVSKYEYDESLIPSPATETSDNISPVASP 906
QY 807 GVDFSITQFVRNLGHEHLMDFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGYVE- 865
DB 907 VHTGFLVSFM-----VDARGGSMRSGRHLGRVVVP 937
QY 866 -----RLISGOGLNPYLTLNLTSGSTLIDLSP----- 894
DB 938 PRTCAAPTRITCRULVKOKLPTPPPLAEEGLASRVIALGPTGVQFLSPVIVEIPHFASQ 997
QY 895 --DKREFOSVEEMOSTVREHRRDGGHAGGIFNRYNLIKIQKVCN 936
DB 998 GRGDRELVLRSENGSVNKEH-----NRYGESYLDQILN 1032
RESULT 7
ID 070511 PRELIMINARY; PRT: 2622 AA.
AC 070511;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1999 (TReMBLrel. 11, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).
GN ANK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98417633; PubMed=9744885;
RA Zhang X., Bennett V.;
RT "Restriction of 480/270-kD ankyrin G to axon proximal segments
RT requires multiple ankyrin G-specific domains.";
RJ J. Cell Biol. 142:1571-1581(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Carpenter S.S., Zhang X.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 934-1220 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;
RX MEDLINE=98330444; PubMed=9664041;
RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RT "Ankyrin G is associated with the postsynaptic membrane and the
RT sarcoplasmic reticulum in the skeletal muscle fibre.";
RJ J. Cell Sci. 111:2197-2207(1998).
DR EMBL; AF102552; AAC78143.1; -
DR EMBL; AF065150; AAC18853.1; -
DR HSP; P5273; 1B18.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR00488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
FT NON_TER 2622 2622
SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

```

Query Match 15.0%; Score 864.5; DB 11; Length 2622;  
 Best Local Similarity 30.5%; Pred. No. 4.8e-46;  
 Matches 276; Conservative 123; Mismatches 360; Indels 147; Gaps 23;

```

QY 3 GRKDVVEYLQNGASVQARDGGGLIPLHNACSFHAEVYNLLLRHGCADPNARDNNMYTPL 62
DB 85 GHVEVVELQREANVDAATKKGNALHTASLAGOAEVVKVLTNCANNAQSQNGFTPL 144
QY 63 HBAALTKGIDVICIVILHQGAETIRNTDGTALDLA-----DPSAKAVLTGEYK---KDE 114
DB 145 YMAAQENHLEVYRFLIDNGASQSLATEDGFTPLAVALOQGHQVLSLLLENDTKGVRPLP 204
QY 115 LLESARSGNEEKMMALL-----TPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLL 164
DB 205 ALHIAARKDDTTAAALLLQNDTINADIESKMVNVNRATESG--FTSLHIAAHYGNINVTALL 262
QY 165 LOHGADVHAKDGLVPLHNACSYGHEVTELLVHKGCANVMDLWQFTPLHEAASKNRV 224
DB 263 LNRAAAVDFTANDITPLHVASKRGNANVKKLLDRGAKIDAKTRDGLTPLHCGARSHE 322
QY 225 EYCSLLLSYGADPTLLNCHINKSAIDLAPTPOLKERLAYEFGHSLLOAAAREADVTIKKH 284
DB 323 QVEMLLDRAA--PILSKTKN-----GLSPUHMATQGD-----H 354
QY 285 LSEMVNFKH-----POTHETALHCAAAPYPRKRKQICEILLRKGANINEKTKFLTP 337
DB 355 LNCVOLLLOHNVVDVTDNYLTALHVAHCHGYK---VAKVLLDKKANPNKALNGFTP 411
QY 338 LHVASEKAHNDVVEVVKHEAKVNALDNIGQTSLHRAAYCGHLOTCLRLLSYGCDDNIIIS 397
DB 412 LHIACKNRIRVWELLKKGASTQAVTESGLTPIHVAAFMGHVNIVSQLMHGASPNNTN 471
QY 398 LOGFTALQMG-----NENYQOLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-Q 449
DB 472 VRGETALHMAARSQGAQEVVRYLVODGAQVEAKAKDDPTPLHISARLGKADIVQOLLQOGA 531
QY 450 SYNCRDIEGRQSTPLHFAAGYNRVSVVEYLLOHGADVHAKDKGLVPLHNACSYGHVEVA 509
DB 532 SPNAATTSG--YTPHLHSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAYGKLEVA 589
QY 510 ELLVHKGVVNVADLWKFTPLHEAAAKGYETICKLLQHGADPTKKNRQGNTPDL-LVKD 568
DB 590 SULLQKASPDAAAGKSGLTPLHVAHYDNQKVALLLLDQOCASPHAAAKNGYTPLHIAAKK 649
QY 569 GDTDIQDLL-----RGDAALLDAKKGCLARVKKLSSPD--NVNCRDTQGRHST 615
DB 650 NQMDJATSLLEYGADANPVYTRGIAVSHLAAQEGHVDVMSLLSRNANVNLNKGSG--LT 707
QY 616 PLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDAALIKYNACVYNAT 675
DB 708 PLHLAAGQEDRVNVAEVLVNOGAHVDAQTKMGYTPLVHVGCHYGNIKIVNFLQHSKVNAK 767
QY 676 DKWAFPLHEAAQKGRQTOLCALLAHGADPTLKNQGGQTPDLVS-----ADDSALL 728
DB 768 TKNGYTPLHQAQOQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYSVYVDTLKVVT 827
QY 729 TAAMPSPALPSCYK---POVLNGVRSRPGATADALSSGSPSSLSAASLDNLS-GSFE 784
DB 828 EEIMTTTITERKHKNVPEPMNEVL-----MSDDEVGKASAEKLSDEY-- 873
QY 785 LSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHK 844
DB 874 -----ISDGEEG-----EDAITGDTDKYLGQP 895
QY 845 ELKEIG 850
DB 896 DLKELG 901
RESULT 8
ID 088521 PRELIMINARY; PRT: 1762 AA.
AC 088521;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE 190 kDa ankyrin isoform.

```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395113; PubMed=9727010;
RA Thevananthar S., Kollu A.H., Devarajan P.;
RT "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT lung that associates with the plasma membrane and binds alpha-Na, k-
RT ATPase.";
RL J. Biol. Chem. 273:23952-23958(1998).
DR EMBL; AF069525; AAC34809.1; -.
DR HSP; P55273; 1B18.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000169; Shprot_acsite.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 19.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1762 AA; 191899 MW; 9023280086A7BF4E CRC64;

Query Match          15.0%; Score 863.5; DB 11; Length 1762;
Best Local Similarity 30.2%; Pred. No. 3.1e-46;
Matches 270; Conservative 123; Mismatches 369; Indels 133; Gaps 21;

Qy 3 GRKDVVEYLLQNGASVQARDGGGLPLHNACSFHAEVNVNLLLRHGADPNARDNNYTP 62
Db 79 GIVEVSELLQREANVDQPTKGNLTALHIAAGVAVKVLVTNGANVNAQSGNFTPL 138
Qy 63 HEATKGIKIDVCIVLQHGAEPTIRNTDGTALDIA-----DPSAKAVLTGEYK---RDE 114
Db 139 YMAAENHLEVVRFLLDNGASGLATEOGFTPLAVALOQGHQDVVSLLENDTKGKVR 198
Qy 115 LESARSGNEKMMALPTPLNVNCHASDGKSTPLHLAAGYNVRKIVOLLQHGADVHAK 174
Db 199 ALHIAARKDDTKAAALLLQNDTNADIESKSGFTPLHIAAHYGNINVTALLNRAAVDT 258
Qy 175 DKGDLVPLHNACSYHYETVELLVKHGACVAMDLWQFTPLHEAASKRVEVCSLLLSYG 234
Db 259 ARNDITPLHVASKRGNNVNNKLLLDRGAKIDAKTRDGLTPLHCGARGSGHEQVVELLDRA 318
Qy 235 ADPTLLNCHNSAIDLAPTQPKERLAYEFKGHSLQAAAREADVTRIKHLSLEMVNFKH 294
Db 319 A-PILSKTN-----GLSLPHMATOGD-----HLNCVOLLQOH 350
Qy 295 -----POTETALHCAASPYPKRQICELLIRKGANINEKTEFTPLHVASEKAHN 347
Db 351 NVPVDVNTDYLTAHVAACHGHYK---VAKVLLDKKANPAKALNGFTPLHIAACKNRI 407
Qy 348 DVEVVVKHAKVNALDNLGOTSLHRAAYCGHLQTCRLLLSYGCDPNLIISLGFTALOMG 407
Db 408 RMELLKHGASIQAVTSRGFTPLHVAAFMGHVNITVSQLMHGASPNNTNVRGETALHMA 467
Qy 408 -----NENVOILLQBSISLGNSEADQ---LLEAAKAGADVETVKK-LCTVQSVNCRDIEGR 459
Db 468 ARSGQAEVVRVLYVDGAQVEAKARDEQHTPHFSRLGAEIVQVQLQOAGSPNAATTSG- 526
Qy 460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEAVALLVKHGAVV 519
Db 527 -YTPLHLARSREGHEVDAAFLDHDGAFSLTITTKGFTPLHVAAYKGLVASLLQKSASP 585
Qy 520 NVADLWKFPTPLHEAAKKGKYEICKLLLLQHGADPTKKNRDNTPDLDL-VKDGDTDIQDILL- 577

```

```

Db 586 DAAGKSGLPFLHVAHYDNQKVALLLDQASPHAAKNGYTPPLHIAAKNQMDIATSL 645
Qy 578 -----RGDAALLDAKKGGLARVKKLSSPD-NVNCRDTCGRHSTPLHLAAGYNN 625
Db 646 EYGADANAVTROGIASVHLAAQEGHVDVSLLSRANVNLSNKRGLN--PLHLGGQEDR 703
Qy 626 LEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHE 685
Db 704 VNVAEVLYNQGHAVDAQTKMGYTPLHVGCHYGNIKIVNELLQHSKAVNAKTNGYTP 763
Qy 686 AAKGRTQLCALLHAGADPTLKNQEGQTPDLVS-----ADDVSALLTAAMPSPSALP 738
Db 764 AAQOQGHTHIINVLLQNNASPNELTVNGNTALAIARLGVISVVDTLKVVTEELMTIT 823
Qy 739 SCYK---POVLNGVRSPGATADALSGSPSSLSAASLDNLSGSFSELSSVSSSGT 795
Db 824 EXHKMNVPEMNEVLVD--MSDDEVGKG-SAPEKLSHGEYI-----SDGEE 865
Qy 796 GASLLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 850
Db 866 G-----EDAITGDTDKSLGPODLKELG 887

RESULT 9
Q13768 PRELIMINARY; PRT; 1719 AA.
AC Q13768;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Alt. ankyrin (variant 2.2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEMATOPOIETIC;
RX MEDLINE=90158830; PubMed=2137557;
RA Lux S.E., John K.M., Bennett V.;
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT structure with homology to tissue-differentiation and cell-cycle
RT control proteins.";
RL Nature 344:36-42(1990).
DR EMBL; X16609; CAA34611.1; -.
DR HSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1719 AA; 189011 MW; F63465D16D975CBF CRC64;

```

```

Query Match          14.6%; Score 842; DB 4; Length 1719;
Best Local Similarity 29.9%; Pred. No. 7e-45;
Matches 240; Conservative 122; Mismatches 264; Indels 177; Gaps 20;

Qy 3 GRKDVVEYLLQNGASVQARDGGGLPLHNACSFHAEVNVNLLLRHGADPNARDNNYTP 62
Db 89 QDDEVVRELVNKGANVNAQSKGFTPLVMAAQENHLEVVVKFLENGANQVATEDGFTPL 148
Qy 63 HEATKGIKIDVCIVLQHGAEPTI-----RNTDGTALDLADPSAKAVLTGEYK 113

```

```

Db 149 AVALQOQHENVVAHLINYGTKGVRLPALHIAARNDDTRTA-----AVLQNDPNP 199
QY 114 ELESARSNGEEMKMLLTPLNVNCHASDGRKS-----TPHLAAGY 155
Db 200 DVL--SKTG-----FTPLHIAAHYENLVQAQLLNRGASVNFTPONGITPLHIASRR 249
QY 156 NRKIVOLLQHGADYHAKDGLVPLHNACSYGHYEVELLKVHKGACVNMALDQWFTPL 215
Db 250 GNVIMVRLLDGQAQIETKTKDELTPHCAARNGHVRISIELLDHGAPIAKTKNGLSPI 309
QY 216 HEARKNRVCSLLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKHSLLQARE 275
Db 310 HMAAQGHDLQVRLLLQYDAE-----ID----- 332
QY 276 ADVTRIKHLSLEMVNEKHPQTHETALHCAASPYPKRKOICELLRKGANINEKTKREFL 335
Db 333 -DIT-----LDHLTPLHVA-----HCCHHVRVAKVLLDKGAKPNSRALNGF 372
QY 336 TPLHVASEKAHNDVVEVVVHKEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNI 395
Db 373 TPLHIACKKNHVRVMEILLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLQORGASPNV 432
QY 396 ISLOGFTALQW-----NENVQQLQOE-----GISLGNSEADROLLE- 432
Db 433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPHCAARIGHTNMVYKLLLEN 492
QY 433 -----AAKAGDVETVKKLTQVQ-SYNCRDIEGROSTPLHFAAGYNRVS 474
Db 493 NANPNLATTAGHTPLHIAAREGHVETVALLERKASQACMTKKG--FTPLHVAAYKQVR 550
QY 475 VVEYLLOHGADVHAKDGGVPLHNACSYGHYEVAELLVKHGAVNVNADLWK-FTPLHEA 533
Db 551 VAELLERDAHPNAAGKNGLTPLHVAHNNLDIVKLLPRGGSPPH-SPAWNGYTPLHIA 609
QY 534 AAKGYEICKLLOHGADPTKKNRDGNTPLDL-VKGDGTDIQLDRLGDAALLDAKKGCL 592
Db 610 AKQNOVEVARSLQYGGSSANAESVQGVTPHLHAAQEGHAEMVALL----- 654
QY 593 ARVKKLSPPNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAODKGGILPLHN 652
Db 655 -----LSQANGNLGNKSG--LTPHLHVAQEGHPVADVLIKGVWVDATTRMGYTPLHV 707
QY 653 AASYGHVDVVAALLIKYNACVNATDKWAFTPLHEAAQKRGRTQLCALLAHGADPTLKNQEG 712
Db 708 ASHYGNKIKVFLQHQADVNAKTKGYSPLHQAQOQGHDTIVTLLKNGASPNVSSDG 767
QY 713 QTPLDL-----VSADDVSALLT 729
Db 768 TTPHIAIAKRLGYISVTDVLKVV 790

RESULT 10
P97582
ID P97582 PRELIMINARY; PRT; 843 AA.
AC P97582;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ankyrin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RA Cosentino M.T., Jones O.T.;
RT "Rat brain ankyrin membrane binding domain."
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U65916; ABA7551.1;
DR HSSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 24.

```

```

DR SMART; SM00248; ANK; 21.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat; Repeat.
FT NON_TER 1 843
SQ SEQUENCE 843 AA; 89982 MW; DC5A6AC78825D223 CRC64;

Query Match 14.6%; Score 841.5; DB 11; Length 843;
Best Local Similarity 29.7%; Pred. No. 2.6e-45;
Matches 265; Conservative 138; Mismatches 331; Indels 137; Gaps 21;

QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSFHGAHYVNNLLRHGADPNARDNMNYTPL 62
Db 40 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLKEGANINAQSQNGFTPL 99
QY 63 HEAAIKKIDVICIVILQHGAEPTIRNTDGR-----ALDADPSAKAVLTGEYKDEL--- 115
Db 100 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQOQHNOAVAILLENDTKGKVRP 159
QY 116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----TPHLAAGYNRKIVQL 163
Db 160 ALHIAARKDDTKSAALLQ-----NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 216
QY 164 LLOHGADVHAKDGLVPLHNACSYGHYEVELLVKHGACVNMALDQWFTPLHEAASKNR 223
Db 217 LLNRGAADVFTARNGITPLHVASKRGNTNMVKLLLDRCGQIDAKTRDGLTPLHCAARSH 276
QY 224 VECSLLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKHSLLQAAAREADVTRKK 283
Db 277 DOVVELLERGA-PLLARTKN-----GLSPLHMAAQGDHVECVK 314
QY 284 HLSLEMVNFKHP-----QTHETALHCAASPYPKRKOICELLRKGANINEKTKREFTPL 338
Db 315 HL-----LOHKAPVDVTDLYDTALHVA-----HCGHYRVTKLLDKRANPNARALNGFTPL 367
QY 339 HVASEKAHNDVVEVVVHKEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDPNIISL 398
Db 368 HIACKKNRIKVMELLVKYGAVIAITTESGLTPIPVAAFMGHLNIVLLLLONGASPDVTNI 427
QY 399 QGFTALQMGNEVQQLQOEGISLGNSEADROLLENAKAGDVETVKKLTQVSVNCRDIEG 458
Db 428 RGETALHN-----AARAGEVEVVR--CLLRNGALVDARA 459
QY 459 R-QSTPLHFAAGYNRVSVEVYLLQHGADVHAKDGGVPLHNACSYGHYEVAELLVKHGA 517
Db 460 REEQTPHLIASRLGKTEIVQLLQHMHPDAATTNGYTPHLISAREGQVDVASVLEAGA 519
QY 518 VVNVDLWKFTPLHEAAAKGYEICKLLOHGADPTKKNRDGNTPLDLVKDGDTDIQLDL 577
Db 520 AHSLATKKGFTPLHVAAYKYSGLDVAKLLQRRRAADSAGKNGLTPLHVAAHYDN----- 573
QY 578 RGDAAALLDAKKGGLARVKKLSPPNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLQHGA 637
Db 574 -----OKVALLLEKGAEP-----HATAKNGYTPHIAAKKNQMOQTASTLLNTGA 618
QY 638 DVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKRGRTQLCAL 697
Db 619 ETNIVTKQGVTPHLHSAQEGHTDWTVLVEKGANIHMTKSGTSLHIAAEEEDKYNVADI 678
QY 698 LLAHGADPTLKNQEGQTPDLVL-----SADDVSALLTAAMPSPA-LPSCYKP----- 743
Db 679 LTKHGADQDAYTKLGYTPLIVACHYGNVKNMNVNLLKOGANVNAKTKNGYTPHQAQOQH 738
QY 744 -QVLNGVRSPGATADALSSGSPSSLSAASSLDNLGSFSELSVSVSSSGTEGASSLEK 802
Db 739 THIIINVLLQHGAKPNATTA--NGNTALAIKRLGYI--SVVDTLKVVTEEVTTTTTITE 794
QY 803 KEVFGVDFSLITQFVRNLGLEHLMIDIFERE--QITLDVLVEMGHKELKEIG 850
Db 795 KHLNAPETMTE-----VLQVSDDEGDDTVTGDGGEYLRPDLRELK 836

```





RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003845; AAF59369.2; -.  
DR HSSP: P80144; 2MYO.  
DR FlyBase: FBgn0011747; Ank.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000488; Death.  
DR InterPro: IPR000906; ZUS.  
DR Pfam: PF00023; ank; 23.  
DR Pfam: PF00531; death; 1.  
DR Pfam: PF00791; ZUS; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 21.  
DR SMART: SM00005; DEATH; 1.  
DR SMART: SM00218; ZUS; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 21.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE: PS50017; DEATH\_DOMAIN; 1.  
KW ANK repeat: Repeat.  
SQ SEQUENCE 1549 AA; 170167 MW; 3ECEC852B1484DFF CRC64;  
  
Query Match 14.5%; Score 836; DB 5; Length 1549;  
Best Local Similarity 27.4%; Pred. No. 1.4e-44;  
Matches 298; Conservative 166; Mismatches 469; Indels 154; Gaps 32;  
  
QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSGHAEVNVLLRHGADPNARDNNWYTPL 62  
DB 84 GYVDICCELLRRGIKIDNATKKGNTALHTIASLAGQHDVINQLILYNNANVQSLNGFTPL 143  
  
QY 63 HEAAIKGIDVCIVLQHGAEPTIRNTDGR TALDLA-----DPSAKAVLTGEYK---KDE 114  
DB 144 YMAAQENHDMCCRTLLANGANSLSSTEDGFTPLAVAMQOCHDKIVAVLENDVRGKVRLP 203  
  
QY 115 LLESARGNEKMMALITPLNVNCHASDGRKSTPLHLAGYNRVKIVQLLQHGADVHAK 174  
DB 204 ALHIAAKNDVNAKLLQLDHPNADIVSKSGFTPLHIAAHYGNVDIATLLNNKADVNYV 263  
  
QY 175 DKGDVPLHNACSYGHEYTELLVKIGACVNMDLQWFTPLHEAASKNRYEVCSLLSYG 234  
DB 264 AKHNITPLHVACKMGKLSLCTLLCLRGAKIDAAATROGLTPLHCASRSGHVEYIKHLQQN 323  
  
QY 235 ADPTLLNCHN-KSAIDLAPTQPKLRLAYEFKGHLSLLQAAREADVTRIKKHSLE----- 288  
DB 324 A-PILFTKTKNLSALHMAAQE-----HDEAAHLLLDNKPVDVTVLYLTALHVAHC 376  
  
QY 289 -----MVNFK-HPQTHE-----TALHCAASPYPKRKOICELLRLKGANINEKTKFPL 335  
DB 377 GHVKVAKLLIDYKANPNARALNGFTPLHITACKK--NRIMKVELLIKHGANIGATTESGL 433  
  
QY 336 TPLHVASERAHNDVVEVVKHKAVALDNIGQTSFHRAAYCGHLOTCLLLSYGCDPNI 395  
DB 434 TPLHVASFGCINVIYLLQHEASDLPTIRGTPPLHAAANQADITIRILL-SAKVDA 492  
  
QY 396 ISLQGF TAL-----QMGNEN-VQOLLQEGISLGNSEADR-QLLLEAAKAGDVETVKLCTV 448  
DB 493 IAREGOTPLHVASRLGNINIIMLLQHGAFINQAQNDKYSALHIAKEGQENIVQVLLN 552  
  
QY 449 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEY 508  
DB 553 GAEN-NAVTKKGFTPLHLACKYQKQNVQVILLQNGASIDFQGNVDVTPHLHVATHYNNPSI 611  
  
QY 509 AELLVKHGA VNVVADLWKFTPLHEAAKGYEICKLLQLQHGADPTKKNRDNPTPLD-LVK 567  
DB 612 VELLKNGSPNLCARNGOCAIHACKKNYLETAMQLQHGADVNTIISGFSPLHAAQ 671  
  
QY 568 DGDPTDQLLRGDAALLDAAKKGLARVKKLSLPDNNVNCRDTCGRHSTPLHLAGAGYNNLE 627  
DB 672 GGNVDVQVLLLELGVISAAKNGL-----TPLHVAAQEGHVL 708

QY 628 VABYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVYNATDKWAFPTLHEAA 687  
DB 709 VSQILLEHGANISERTNGYTPPLHMAHYCHLDLVKFFIENDADADIEMSSNIGTYTPLHQA 768  
  
QY 688 QKQRTQCALLLAHGADPTLKNOEGQTPDLVSDADDVSALLTAAMPSPALPSYCKKQVPLN 747  
DB 769 QOQHIMIINLLRHKANPNALTCKDNTALHASN-----LGXYVTMESLKIVTSTSVIN 822  
  
QY 748 G-----VRSPGATADALSSGSPSSLSAAS-----SLDNLGSFSELSVWS 790  
DB 823 SNIGATIEELKVMTPPELMQETLLSDSDDECDLLDHNHYKYMATDDLKANYGQDQKNFD 882  
  
QY 791 SSQTE-----GASSLEKKEVPGVDFSIQFVRNLGLEHMD--IFERQITLDLVEMGHK 844  
DB 883 TTNTDHDLTDSVLNKKIELPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSP--- 936  
  
QY 845 ELKEIGINAYGHRHKLKIGVE-----RLISGOOGLNPYLTLTSSGTLID 891  
DB 937 LVDARGGSMRGYRHNGVRIIVPPKACAEPTITCRVYKPORVNVNPPPLMGEALVSRILE 996  
  
QY 892 LSPDDKEFOS---VEEMQSTVREHRDGGHAGGIFNRYNLIKQKVCNKKLWERYTHRRK 948  
DB 997 MSPVDCGMFLSPITLEVPHYGTLRK-----NEREII-ILRSDNGESWREH-NLYK 1043  
  
QY 949 EYSEENHNHANERMLPHGSPFYVNAIIHKGFDERHAYIGMGFAGIYPAENSSKSNQVYVG 1008  
DB 1044 DIIGEDINQTEE---FHSDDRIVR-IVTQNVPHFFAVV-----SRVRQEVHV 1085  
  
QY 1009 IG--GGT 1013  
DB 1086 IGPDDGT 1092  
  
RESULT 14  
Q24241  
ID Q24241 PRELIMINARY; PRT; 1549 AA.  
AC Q24241;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE ANKYRIN  
GN ANK OR CG1651.  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A631;  
RX MEDLINE=95024098; PubMed=7937942;  
RA Dubreuil R.R., Yu J.;  
RT "Ankyrin and beta-spectrin accumulate independently of alpha-spectrin  
in *Drosophila*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:10285-10289(1994).  
DR EMBL: L35601; AAC37208.1; -.  
DR HSSP: P80144; 2MYO.  
DR FlyBase: FBgn0011747; Ank.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000488; Death.  
DR InterPro: IPR000906; ZUS.  
DR Pfam: PF00023; ank; 23.  
DR Pfam: PF00531; death; 1.  
DR Pfam: PF00791; ZUS; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 21.  
DR SMART: SM00005; DEATH; 1.  
DR SMART: SM00218; ZUS; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 21.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE: PS50017; DEATH\_DOMAIN; 1.  
KW ANK repeat: Repeat.  
SQ SEQUENCE 1549 AA; 170096 MW; 587FB6BF88B657CE CRC64;



```
Query Match 14.5%; Score 834; DB 5; Length 1549;
Best Local Similarity 24.8%; Pred. No. 1.9e-44;
Matches 288; Conservative 175; Mismatches 357; Indels 300; Gaps 32;

QY 3 GRKDVVEYLLQNGASVQARDGGLIPLHNACSFHAEVYVNNLLRHGADPNARDNNWYTPL 62
DB 84 GYVDICCELLRRGIDKATKNGTALHIASLAGOHDIQNLILYNNVNVQSLNGETPL 143
QY 63 HEAAIKGKIDVICVILQHGAEPTIRNTDGRALDADPSAKAVLTGEYKDELLESARG 122
DB 144 YMAAENHDNCCRTLLANGANSLSSTEDGFT-----PLAVAMQOG----- 183
QY 123 NEEKMMALLPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLQHGADVHAKDKGLVP 181
DB 184 -HDKIVAVLLENDV-----GKVRPLPALHIAAKNDVNAKLLLOHDPNADIVSKSGTTP 237
QY 182 LHNACSYGHYEVELLVKHGACVNAQMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
DB 238 LHIAAHYGNVDIATLLNNKADVNTYVAKHNITPLHVACKWGKLSLCTLLLCRGA----- 291
QY 242 CHNKSADLAPTPOLKERLAYEFKHSLLQAREADVTIRIKHLSLEMVNFKHQTHETA 301
DB 292 -----KIDAATRDGL-----TP 303
QY 302 LHCAASPYPKRKQICELLRRKGANINEKTEFLPLHVASEKAHNDVVEVVKHEAKVN 361
DB 304 LHCASRSGH---VEVVKHLQQAAPILTKNGLSALHWAAGHDEAAHLLLDNKAPVD 360
QY 362 ALDNLGQTSLHRAAYCGHQLQTCRLLSYGCDDPNTISLQGTALQNGNEN-----VQOQLQ 416
DB 361 ETVTDYLTALHVAHCGHVAKLLLDYKANPNARALNGFTPLHIAACKNRKIKMVELLIK 420
QY 417 EGISLG-----NSEADROLLE-----AAKAGDVET 441
DB 421 HGANIGATTESGLTPLHVASFNGCINVIYLLQHEASADLPTIRGETPLHIAARANQADI 480
QY 442 VKKLCTVQSVNGRDIEGROSTPLHFAAGYNRVSVVEYVILQHGADVHAKD----- 490
DB 481 IRIILRSKAVDAIVREG--TPLHVASRLGNINIIMLLQHGAEINAQSNOKYSALHIAA 538
QY 491 -----KGLVPLHNACSYGHVEAEVALLVKGAVNVNADLWKFPT 528
DB 539 KEGQENIVQVLENGAENNAVTKKGTFTPLHACKYQKQNVVQIILQNGASIDFQCKNDVT 598
QY 529 PLHEA-----AAKGY-EICKLILQHGADPTKK 555
DB 599 PLHVATHYNNPSIVELLLKNGSSPNLCARNGOCAIHIACKKNYLEIAMQLLOHGADVNI 658
QY 556 NRDGNTPLDL-VKDGDDTDIQLLRGDAALLDAKKGLARVKKLSSPDNVNCRDTPGRHS 614
DB 659 SKSGFSPHLAAQGGNVDMVQLLEVGVISAAKNGL----- 695
QY 615 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNACSYGHVDVVAALLIKYNACVNA 674
DB 696 TPLHVAQCGHVLVSQILLEGHANTISERTNGYTPPLHMAAHYGHLDLVKFFTEINDADTEM 755
QY 675 TDKWAFTPLHEAAOKRGTOLCALLAHGADPTLKNQEGOTPLDLSADDSVALLTAAMP 734
DB 756 SSNIGTTPHQAQOQHIMIINLLRHKANPNALTKDGNALHASN-----LGYVTVM 809
QY 735 SALPSCYKQVLLNG-----VRSFGATADALSSGFPSSPSLSAAS-----SLDN 777
DB 810 ESLKIVTSTSVINSNIGATEEKLKVMTPELMQETLLSDSDDESCDDLLDHNHYKYMATDD 869
QY 778 LGSFSELSVSSSGSTE-----GASLEKKEVPDGVDFSIQVRLNGLEHLM-----IFERE 831
DB 870 LKANYQDQKNFDTTNDTDLTDVSLNKKEILPNEMSCIELTE-----ICHKPDNVVIARS 926
QY 832 QITLDVLVEMGHKEKEICINAYGHRHKLKIGVE-----RLISQOQGLNYPYL 878
DB 927 QVHLGLVLSF---LVDARGSGRMGRVNHGVRITVPPKACAEPTTRITCRYVKPQVRVNNPPP 983
```





treating or preventing cell cycle associated disorders such as cancer

Claim 22; Fig 4; 63pp; English.

This sequence is the Tankyrase homologue isotype 2 (TaHo-2) protein of the invention. The invention also relates to the TaHo-2 protein. The TaHo proteins are useful for inducing or preventing cell proliferation in cells, and in the study or treatment of conditions mediated by the cell cycle proteins, such as to diagnose, treat or prevent cell cycle associated disorders, preferably cancer. The TaHo coding sequences are useful as hybridisation probes, in chromosome and gene mapping and in the generation of anti-sense DNA and RNA. The coding sequences are also useful for the preparation of TaHo, for generating either transgenic animals or knock out animals which, in turn, are useful in a development and screening of therapeutically useful agents, in gene therapy, as vaccine, and for construction of hybridisation probes for mapping the gene which encodes TaHo and for the genetic analysis of individuals with genetic disorders. The TaHo proteins, and their coding sequences are useful in screening assays.

Query Match	100.08;	Score 6464;	DB 22;	Length 1240;	
Best Local Similarity	100.08;	Pred. No. 0;			
Matches 1240;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	RC SARAGAGGOGARVGAAGTADPPVTAGSQAARALASASSPGGLALLAGPGLLL	60		
Db	1	RC SARAGAGGOGARVGAAGTADPPVTAGSQAARALASASSPGGLALLAGPGLLL	60		
Qy	61	RL LALLAAVAARIMSGRCAGGAGCAACAAAEVPAARELFEACRNGDVERVRLVTP	120		
Db	61	RL LALLAAVAARIMSGRCAGGAGCAACAAAEVPAARELFEACRNGDVERVRLVTP	120		
Qy	121	EK VNSRD TAGRSTPLHFAAGFRKDVVEYLQNGANVOARDGGLIPLHNACSFGEAEV	180		
Db	121	EK VNSRD TAGRSTPLHFAAGFRKDVVEYLQNGANVOARDGGLIPLHNACSFGEAEV	180		
Qy	181	VN LLRHGADPNARNWNVTPLHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLADP	240		
Db	181	VN LLRHGADPNARNWNVTPLHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLADP	240		
Qy	241	SA KAVLTGEYKKDELLESARSNEEKMMALLTPLNVNCHASDGKSTPLHLAAGYNRVK1	300		
Db	241	SA KAVLTGEYKKDELLESARSNEEKMMALLTPLNVNCHASDGKSTPLHLAAGYNRVK1	300		
Qy	301	VO LLLQHGADVIAKDKGDLVPLHNACSYGHEVETELLYKHGACVAMDLWQFTPLHEAAS	360		
Db	301	VO LLLQHGADVIAKDKGDLVPLHNACSYGHEVETELLYKHGACVAMDLWQFTPLHEAAS	360		
Qy	361	KN RVCSLLLSYGADPTLLNCHNKSADLAPTPO LKERLAYEFKGHSLQQAAREADVTR	420		
Db	361	KN RVCSLLLSYGADPTLLNCHNKSADLAPTPO LKERLAYEFKGHSLQQAAREADVTR	420		
Qy	421	IK KHL SLEMVNFKHPQTHETALHCAAAAPYKPKQICE LLRKGANIN EKTKEFTPLHV	480		
Db	421	IK KHL SLEMVNFKHPQTHETALHCAAAAPYKPKQICE LLRKGANIN EKTKEFTPLHV	480		
Qy	481	ASE KAHNDVVEVVKHEAKVNDLNGQTS LHRAYCGHLQTCRLLLSYCGDPNIIISLOG	540		
Db	481	ASE KAHNDVVEVVKHEAKVNDLNGQTS LHRAYCGHLQTCRLLLSYCGDPNIIISLOG	540		
Qy	541	FTALQMGNEVVO LLQEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGRQ	600		
Db	541	FTALQMGNEVVO LLQEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSVNCRDIEGRQ	600		
Qy	601	STPLHFAAGYNRVSVVEYLLQHGADVIAKDKGGLVPLHNACSYGHEVAE LLVYKHGAVYN	660		
Db	601	STPLHFAAGYNRVSVVEYLLQHGADVIAKDKGGLVPLHNACSYGHEVAE LLVYKHGAVYN	660		
Qy	661	VADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPDLIVKDGDTDIQDLLRGD	720		

Db	661	VADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPDLIVKDGDTDIQDLLRGD	720		
Qy	721	AALIDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV	780		
Db	721	AALIDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV	780		
Qy	781	AODKGGLLPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAOKGRTOQLCALLIA	840		
Db	781	AODKGGLLPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAOKGRTOQLCALLIA	840		
Qy	841	HGADPTLNQEGQTPLDLVSADDDVSALLTAAMPSPALPSCYKPOVLNGVRSFGATADALS	900		
Db	841	HGADPTLNQEGQTPLDLVSADDDVSALLTAAMPSPALPSCYKPOVLNGVRSFGATADALS	900		
Qy	901	SGPSSPSSLSAASSLDNLGSGFSELSSVSSSGTSGASSLEKKEVPGVDFSTIQFVRNIG	960		
Db	901	SGPSSPSSLSAASSLDNLGSGFSELSSVSSSGTSGASSLEKKEVPGVDFSTIQFVRNIG	960		
Qy	961	LEHMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGGQGLNPYLTL	1020		
Db	961	LEHMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGGQGLNPYLTL	1020		
Qy	1021	NTSGSCTILIDLSDDKKEFOSVEEEMQSTVREHROGGHAGGINRYNLIKIOKVCNKKLW	1080		
Db	1021	NTSGSCTILIDLSDDKKEFOSVEEEMQSTVREHROGGHAGGINRYNLIKIOKVCNKKLW	1080		
Qy	1081	ERYTHRRKEVSEENHANERMLFHGSPFVNAI IHKGFDERHAYIGGMFGAGIYFAENSS	1140		
Db	1081	ERYTHRRKEVSEENHANERMLFHGSPFVNAI IHKGFDERHAYIGGMFGAGIYFAENSS	1140		
Qy	1141	KSNOYVYIGGGTGPCVHKDRSCYICHRQLLCFRTVTLGKSFLOFSAMKMAHSPPGHHSVT	1200		
Db	1141	KSNOYVYIGGGTGPCVHKDRSCYICHRQLLCFRTVTLGKSFLOFSAMKMAHSPPGHHSVT	1200		
Qy	1201	GRPSVNGIALAEYVIYRGEQAYPEYLITYQIMRPEGWVDG	1240		
Db	1201	GRPSVNGIALAEYVIYRGEQAYPEYLITYQIMRPEGWVDG	1240		
RESULT 2					
AAB66290					
ID	AAB66290 standard; Protein; 1262 AA.				
XX	AAB66290;				
AC	05-APR-2001 (first entry)				
DT	Human tankyrase2 clone consensus protein SEQ ID NO: 107.				
XX	Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging; inflammatory disorder.				
DE	Homo sapiens.				
XX	WO200100849-A1.				
PN	04-JAN-2001.				
XX	28-JUN-2000; 2000WO-US17827.				
PF	29-JUN-1999; 99US-0141582.				
XX	(ICOS-) ICOS CORP.				
PA	Christenson E, Demaggio AJ, Goldman PS, McElligott DL;				
PI	WPI: 2001-102896/11.				
XX	N-PSDB: AAF63930.				
DR	New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -				
PT					
PT					
XX					

PS Example 2: Page 173-176; 242pp; English.

CC The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.

XX SQ Sequence 1262 AA;

Query Match 98.6%; Score 6375.5; DB 22; Length 1262;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1225; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 6 RGAAGGCGAARGVAAAGTAPDPVTAGQAARALSSAPGGALLLAGPGLLRLLAL 65  
DB 29 RGSRGAGSPARGAR-GRGHGTAPDPVTAGQAARALSSAPGGALLLAGPGLLRLLAL 87  
QY 66 LLAVAAARIMSGRCAGGACASAAAEVPEPAARELFACRNGDVERKRLVTPKVN 125  
DB 88 LLAVAAARIMSGRCAGGACASAAAEVPEPAARELFACRNGDVERKRLVTPKVN 147  
QY 126 RDTAGRKSTPLHFAAGFGKRDVVYLLQNGANVQARDGGILPLHNACSGHAEWNLL 185  
DB 148 RDTAGRKSTPLHFAAGFGKRDVVYLLQNGANVQARDGGILPLHNACSGHAEWNLL 207  
QY 186 RHGADPNARDNWTPLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAV 245  
DB 208 RHGADPNARDNWTPLHEAAIKGIDVCIVLLQHGAEPTIRNTDGRALDADPSAKAV 267  
QY 246 LTGEYKDELESARSNEEKMMALLPLNVNCHASDGRKSTPLHLAGYNRKIVQLLL 305  
DB 268 LTGEYKDELESARSNEEKMMALLPLNVNCHASDGRKSTPLHLAGYNRKIVQLLL 327  
QY 306 QHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNAWDLWQFTPLHEAASKNRVE 365  
DB 328 QHGADVHAKDKGDLVPLHNACSYGHEVTELLVKGACVNAWDLWQFTPLHEAASKNRVE 387  
QY 366 VCSLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKSHSLQAAAREADVTRIKKHL 425  
DB 388 VCSLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKSHSLQAAAREADVTRIKKHL 447  
QY 426 SLEMVNFKHPQTHETALHCAASAPYPRKQICEILLRKGANINEKTEFTPLHVASEKA 485  
DB 448 SLEMVNFKHPQTHETALHCAASAPYPRKQICEILLRKGANINEKTEFTPLHVASEKA 507  
QY 486 HNDVVEVVKHAEKVNALDNLGQTSLHRAAYCGHLOQTCRLLLSYGCDDPNIISLQGFALQ 545  
DB 508 HNDVVEVVKHAEKVNALDNLGQTSLHRAAYCGHLOQTCRLLLSYGCDDPNIISLQGFALQ 567  
QY 546 MGENVQOOLQEGISLQNSADROLLBAKAGDVETVKLCITQSVNCRDIEGRQSTPLH 605  
DB 568 MGENVQOOLQEGISLQNSADROLLBAKAGDVETVKLCITQSVNCRDIEGRQSTPLH 627  
QY 606 FAAGYNRVSVVEYLQHGADVHAKDKGGLVPLHNACSYGHEVAEELLVKCAVNVADLW 665  
DB 628 FAAGYNRVSVVEYLQHGADVHAKDKGGLVPLHNACSYGHEVAEELLVKCAVNVADLW 687  
QY 666 KFTPLHEAAAKGYEICKLLQHGADPTKKNRQDNTPLDLVKDGDTDIQLLRGDAALLD 725  
DB 688 KFTPLHEAAAKGYEICKLLQHGADPTKKNRQDNTPLDLVKDGDTDIQLLRGDAALLD 747  
QY 726 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 785  
DB 748 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 807  
QY 786 GLTPLHNAASGYHVDVAALLIKYNACVNAATDKWAFTPLHEAAQKGRTOLCALLHAADP 845  
DB 808 GLTPLHNAASGYHVDVAALLIKYNACVNAATDKWAFTPLHEAAQKGRTOLCALLHAADP 867

QY 846 TLKNOEGOTPLDLVSADDVVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGPSS 905  
DB 868 TLKNOEGOTPLDLVSADDVVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSSGPSS 927  
QY 906 PSSLSAASSLNLGSGFSFSELSSVSSVSSSGTEGASSLEKKEVPGVDFSIQTFVRNLGLEHLM 965  
DB 928 PSSLSAASSLNLGSGFSFSELSSVSSVSSSGTEGASSLEKKEVPGVDFSIQTFVRNLGLEHLM 987  
QY 966 DIFEREQITLDLVEMGHKELKEICINAYGHRHKLKIGVERLISQOQGLNPYLLTNTSGS 1025  
DB 988 DIFEREQITLDLVEMGHKELKEICINAYGHRHKLKIGVERLISQOQGLNPYLLTNTSGS 1047  
QY 1026 GTILDLSPDDKEFQSVSEEMQSTVREHRDGGHAGCIENRYNLIKIQVCKNKKLWERTH 1085  
DB 1048 GTILDLSPDDKEFQSVSEEMQSTVREHRDGGHAGCIENRYNLIKIQVCKNKKLWERTH 1107  
QY 1086 RRKEVSEENHNANERMLFHGSPFVNAIIHKGFDRHAYIGGMFGAGIYFAENSSKSNQY 1145  
DB 1108 RRKEVSEENHNANERMLFHGSPFVNAIIHKGFDRHAYIGGMFGAGIYFAENSSKSNQY 1167  
QY 1146 VYIGGGTGCPCVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSV 1205  
DB 1168 VYIGGGTGCPCVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSV 1227  
QY 1206 NGLALAEVVIYRGQAYPEYLITYQIMRPEGMVDG 1240  
DB 1228 NGLALAEVVIYRGQAYPEYLITYQIMRPEGMVDG 1262

RESULT 3  
AAB66294  
ID AAB66294 standard; Protein: 1385 AA.  
XX AAB66294;  
XX 05-APR-2001 (first entry)  
XX Human tankyrase2 TANK2-LONG SEQ ID NO: 133.  
DE Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
KW Inflammatory disorder.  
OS Homo sapiens.  
XX WO200100849-A1.  
XX 04-JAN-2001.  
XX 28-JUN-2000; 2000WO-US17827.  
XX 29-JUN-1999; 99US-0141582.  
XX (ICOS-) ICOS CORP.  
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
WPI: 2001-102896/11.  
N-PSDB: AAF63952.  
XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
PT inflammatory and autoimmune disorders -  
XX Claim 2; Page 191-194; 242pp; English.  
XX The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.

Db	1111	DIFEREQITLDVLVMGHKELKEIGINAYGHRHKLIKVVERLISGQOGLNPYLTLTNPSGS	117
Qy	1026	GTILIDLSPDDKFEQSVSEEMQSTVREHRDGGHAGGIFNRYNLIKIQKVCNKKLWERYTH	1085
Db	1171	GTILIDLSPDDKFEQSVSEEMQSTVREHRDGGHAGGIFNRYNLIKIQKVCNKKLWERYTH	1230
Qy	1086	RREVESEHNHNERMLFHGSPFYNAIIHKGFDERHAYIGMGFAGIYFAENSSKSNOY	1145
Db	1231	RREVESEHNHNERMLFHGSPFYNAIIHKGFDERHAYIGMGFAGIYFAENSSKSNOY	1290
Qy	1146	VYIGGGTGCPVHKDRSCVICHRRQLLCFVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSV	1205
Db	1291	VYIGGGTGCPVHKDRSCVICHRRQLLCFVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSV	1350
Qy	1206	NGLALAEYVIYRGEQAYPEYLITYOIMRPEGMVDG	1240
Db	1351	NGLALAEYVIYRGEQAYPEYLITYOIMRPEGMVDG	1385
RESULT	4		
AAB66278	ID	AAB66278 standard; Protein; 1169 AA.	
AC	AAB66278;		
XX	05-APR-2001	(first entry)	
DT			
XX			
DE		Human tankyrase2 related protein sequence SEQ ID NO: 2.	
XX			
KW		Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;	
KW		inflammatory disorder.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200100849-A1.	
XX			
PD		04-JAN-2001.	
XX			
PF		28-JUN-2000; 2000WO-US17827.	
XX			
PR		29-JUN-1999; 99US-0141582.	
XX			
PA		(ICOS-) ICOS CORP.	
XX			
PI		Christenson E, Demaggio AJ, Goldman PS, McElligott DL;	
XX			
DR		WPI; 2001-102896/11.	
DR		N-PSDB; AAF63837.	
XX			
PT		New tankyrase2 polypeptides, useful for treating conditions mediated by	
PT		poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,	
PT		inflammatory and autoimmune disorders -	
XX			
PS		Disclosure; Page 109-113; 242pp; English.	
XX			
CC		The present invention provides the protein and coding sequence for the	
CC		human tankyrase2 protein. This is found in two different versions,	
CC		designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has	
CC		polyADP-ribosylation activity and is involved in the modification of	
CC		TRF1, which is a telomere-specific binding protein. The regulation of	
CC		telomere length, in which TRF1 has a role, is linked to ageing and	
CC		cancer. The sequences are useful in the treatment of cancers and	
CC		inflammatory disorders.	
XX			
SQ		Sequence 1169 AA;	
		Query Match 94.7%; Score 6121; DB 22; Length 1169;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps	
Qy	72	ARIMSGRRCAGGACACASAAAEVPAARLFEACRNGDVERKRLVTPKVNRSRDTAGR	131
Db	1	ARIMSGRCAGGACACASAAAEVPAARLFEACRNGDVERKRLVTPKVNRSRDTAGR	60

QY 132 KSTPLHFAAGFRKDVVEYLLQNGANVOARDGGLIPLHNACSGHAEVNNLLRHGADP 191  
Db 61 KSTPLHFAAGFRKDVVEYLLQNGANVOARDGGLIPLHNACSGHAEVNNLLRHGADP 120  
QY 192 NARDNNWYTPHAEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 251  
Db 121 NARDNNWYTPHAEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 180  
QY 252 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 311  
Db 181 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 240  
QY 312 HAKDKGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCISLL 371  
Db 241 HAKDKGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCISLL 300  
QY 372 SYGADPTLLCHNKSAIDLAFTPOLKERLAYEFKGHSLQAAREADVTRIKKHSLEWVN 431  
Db 301 SYGADPTLLCHNKSAIDLAFTPOLKERLAYEFKGHSLQAAREADVTRIKKHSLEWVN 360  
QY 432 FKHPQTHETALHCAASAPYPRKQICELLRLKGANINEKTFEPLTHVASEKAHNDVVE 491  
Db 361 FKHPQTHETALHCAASAPYPRKQICELLRLKGANINEKTFEPLTHVASEKAHNDVVE 420  
QY 492 VVYKHEAKVNALNLGOTSLSHRAAYCGHLOTCTRLLSYGCDDPNTIISLOGFTALQMGNEV 551  
Db 421 VVYKHEAKVNALNLGOTSLSHRAAYCGHLOTCTRLLSYGCDDPNTIISLOGFTALQMGNEV 480  
QY 552 QOLLOEGISLGNSEADROLLEAAKAGDVETVKKICTVQSVNCRDIEGRQSTPLHFAAGYN 611  
Db 481 QOLLOEGISLGNSEADROLLEAAKAGDVETVKKICTVQSVNCRDIEGRQSTPLHFAAGYN 540  
QY 612 RVSVEYLLQHGADVHAKDKGLVPLHNACSYGHEVYAEALLVKGAVNVADLWKFTPLH 671  
Db 541 RVSVEYLLQHGADVHAKDKGLVPLHNACSYGHEVYAEALLVKGAVNVADLWKFTPLH 600  
QY 672 EAAKGYETCKLLQHGADPTKKNRDGNTPLDLVKDGDITDIDLLRGDAALLDAKKGC 731  
Db 601 EAAKGYETCKLLQHGADPTKKNRDGNTPLDLVKDGDITDIDLLRGDAALLDAKKGC 660  
QY 732 LARVKLSSPDNCRDTCGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLH 791  
Db 661 LARVKLSSPDNCRDTCGRHSTPLHLAAGYNLEVAEYLLQHGADVNAQDKGGLIPLH 720  
QY 792 NAASGYHVDVAALLIKYNACVNATDKWAFPLHEAAQKGTQCALLLAHGADPTLKNQE 851  
Db 721 NAASGYHVDVAALLIKYNACVNATDKWAFPLHEAAQKGTQCALLLAHGADPTLKNQE 780  
QY 852 GQTPDLVVSADVSALLTAAMPSPALPSCYKQVNLGVRSPGATADALSSGSPSSLSA 911  
Db 781 GQTPDLVVSADVSALLTAAMPSPALPSCYKQVNLGVRSPGATADALSSGSPSSLSA 840  
QY 912 ASSLDNLGSGFSSELSSVSSSGTEGASSLEKKEVPVDFSIOTFVRNLGLEHLMDFERE 971  
Db 841 ASSLDNLGSGFSSELSSVSSSGTEGASSLEKKEVPVDFSIOTFVRNLGLEHLMDFERE 900  
QY 972 QITLDLVEMGHKELKEIGINAYGHRHKLKGVRLISGOOGLNPYLTLMTSGSGTILID 1031  
Db 901 QITLDLVEMGHKELKEIGINAYGHRHKLKGVRLISGOOGLNPYLTLMTSGSGTILID 960  
QY 1032 LSPDDKEFQSVSEEMQSTVREHRDGGHAGGIFNRYNLTIKQVCNKKLWERYTHRRKEVS 1091  
Db 961 LSPDDKEFQSVSEEMQSTVREHRDGGHAGGIFNRYNLTIKQVCNKKLWERYTHRRKEVS 1020  
QY 1092 EENHNHANERMLPHGSPFNVAIIHKGFDERHAYIGMFGAGIYFAENSSKNQYVYGIG 1151  
Db 1021 EENHNHANERMLPHGSPFNVAIIHKGFDERHAYIGMFGAGIYFAENSSKNQYVYGIG 1080  
QY 1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHSVTRGRPSVNGLALA 1211  
Db 1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGHSVTRGRPSVNGLALA 1140

QY 1212 EYVYRGQAYPEYLITYQIMRPEGVDG 1240  
Db 1141 EYVYRGQAYPEYLITYQIMRPEGVDG 1169

RESULT 5  
AAB66288

XX AAB66288 standard; Protein; 1169 AA.

XX AC AAB66288;

XX XX 05-APR-2001 (first entry)

XX Human tankyrase2 clone consensus protein SEQ ID NO: 101.

XX DE Human: tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
XX KW inflammatory disorder.

XX OS Homo sapiens.

XX PN WO200100849-A1.

XX PD 04-JAN-2001.

XX PF 28-JUN-2000; 2000WO-US17827.

XX PR 29-JUN-1999; 99US-0141582.

XX XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
XX PI N-PSDB; AAF63926.

XX DR WPT: 2001-102896/11.

XX PT New tankyrase2 polypeptides, useful for treating conditions mediated by  
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
PT inflammatory and autoimmune disorders -  
XX XX  
XX Example 1; Page 162-1665; 242pp; English.

XX The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.

XX SQ Sequence 1169 AA;

Query Match 94.7%; Score 6121; DB 22; Length 1169;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 ARMSGRRRCAGGGAACASAAAAEAPAEARELFEACRNGDVERVRLVTPKVNSRDAGR 131  
Db 1 ARMSGRRRCAGGGAACASAAAAEAPAEARELFEACRNGDVERVRLVTPKVNSRDAGR 60

QY 132 KSTPLHFAAGFRKDVVEYLLQNGANVOARDGGLIPLHNACSGHAEVNNLLRHGADP 191  
Db 61 KSTPLHFAAGFRKDVVEYLLQNGANVOARDGGLIPLHNACSGHAEVNNLLRHGADP 120

QY 192 NARDNNWYTPHAEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 251  
Db 121 NARDNNWYTPHAEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYK 180

QY 252 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 311  
Db 181 KDELLESARSNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADV 240

QY 312 HAKDKGLVPLHNACSYGHEVTELLVKGACVNMADLWQFTPLHEAASKNRVEVCISLL 371

```
Db 241 HAKDKGLVPLHNACSYGHYEYETELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLL 300
Qy 372 SYGADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVN 431
Db 301 SYGADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVN 360
Qy 432 FKHPOTHETALHCAASAPYPRKQICELLRKGANINEKTEFLPLHVASEKAHNDVVE 491
Db 361 FKHPOTHETALHCAASAPYPRKQICELLRKGANINEKTEFLPLHVASEKAHNDVVE 420
Qy 492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEV 551
Db 421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEV 480
Qy 552 QQLLEGISLGNSEADROLLEAAKAGDVETVKKLTQVSQVNCRDIEGRQSTPLHFAAGYN 611
Db 481 QQLLEGISLGNSEADROLLEAAKAGDVETVKKLTQVSQVNCRDIEGRQSTPLHFAAGYN 540
Qy 612 RVSVEYLLQHCADVHAKDKGLVPLHNACSYGHYEYETELLVKGACVAMDLWQFTPLH 671
Db 541 RVSVEYLLQHCADVHAKDKGLVPLHNACSYGHYEYETELLVKGACVAMDLWQFTPLH 600
Qy 672 EAAAKGVEICKLLQHGADPTKKNRDGNTPLDLVKDGDITQDLLRGDAALLDAAKKGC 731
Db 601 EAAAKGVEICKLLQHGADPTKKNRDGNTPLDLVKDGDITQDLLRGDAALLDAAKKGC 660
Qy 732 LARVKKLSPDNVNCRDITQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLH 791
Db 661 LARVKKLSPDNVNCRDITQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAODKGLIPLH 720
Qy 792 NAAASYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQICALLAHGADPTLKNOE 851
Db 721 NAAASYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQICALLAHGADPTLKNOE 780
Qy 852 GOTPLDLVSADVSALLTAAMPSPALPSYCKEQVLNGVRSPOGATADALSSGSPSSLSA 911
Db 781 GOTPLDLVSADVSALLTAAMPSPALPSYCKEQVLNGVRSPOGATADALSSGSPSSLSA 840
Qy 912 ASSLDNLSSFSSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFERE 971
Db 841 ASSLDNLSSFSSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFERE 900
Qy 972 QITLDVLVEMGHKELKEIGINAYGHRHKLIGVERLISQOQGLNPYLTINTSGSTILID 1031
Db 901 QITLDVLVEMGHKELKEIGINAYGHRHKLIGVERLISQOQGLNPYLTINTSGSTILID 960
Qy 1032 LSPDDKEFOSVEEEMQSTVREHRDGGHAGGIENRNYNLKIQVCNKKLWERYTHRRKEYS 1091
Db 961 LSPDDKEFOSVEEEMQSTVREHRDGGHAGGIENRNYNLKIQVCNKKLWERYTHRRKEYS 1020
Qy 1092 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGIG 1151
Db 1021 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGMFGAGIYFAENSCKSNQYVYGIG 1080
Qy 1152 GTGCPVHKDRSCYICHRQLFCRVTLGKSFLOFSAMKMAHSPGGHHSVTGRPSVNGLALA 1211
Db 1081 GTGCPVHKDRSCYICHRQLFCRVTLGKSFLOFSAMKMAHSPGGHHSVTGRPSVNGLALA 1140
Qy 1212 EYVYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db 1141 EYVYRGEQAYPEYLITYQIMRPEGMVDG 1169
```

## RESULT 6

AAB66295

ID AAB66295 standard; Protein; 1166 AA.

XX AC

XX AAB66295;

XX DT

XX 05-APR-2001 (first entry)

XX DE

Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.

```
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX Homo sapiens.
OS WO200100849-A1.
PN 04-JAN-2001.
PD 28-JUN-2000; 2000WO-US17827.
PF 29-JUN-1999; 99US-0141582.
PR (ICOS-) ICOS CORP.
PA Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX WPI; 2001-102896/11.
XX N-PSDB; AAF63953.
DR New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders -
XX Claim 3; Page 200-203; 242pp; English.
XX The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
CC polyADP-ribosylation activity and is involved in the modification of
CC TRF1, which is a telomere-specific binding protein. The regulation of
CC telomere length, in which TRF1 has a role, is linked to aging and
CC cancer. The sequences are useful in the treatment of cancers and
CC inflammatory disorders.
XX SQ Sequence 1166 AA;
```

```
Query Match 94.5%; Score 6108; DB 22; Length 1166;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 75 MSGRCAGGACACASAAAEVPAARELFECACNGDVERVKRLVTEKVNRSRTAGRKST 134
Db 1 MSGRCAGGACACASAAAEVPAARELFECACNGDVERVKRLVTEKVNRSRTAGRKST 60
Qy 135 PLHFAAGFRKDVVEYLLQNGANVOARDGGLPLHNACSFHAEVYVNLRLRHGADPNAR 194
Db 61 PLHFAAGFRKDVVEYLLQNGANVOARDGGLPLHNACSFHAEVYVNLRLRHGADPNAR 120
Qy 195 DNWNVTPLHEAAIKGIDYCVILLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 254
Db 121 DNWNVTPLHEAAIKGIDYCVILLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 180
Qy 255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 314
Db 181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 240
Qy 315 DKGDVLPLHNACSYGHYEYETELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLSYG 374
Db 241 DKGDVLPLHNACSYGHYEYETELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLSYG 300
Qy 375 ADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVNFKH 434
Db 301 ADPTLLNCHNKSAIDLAFTPQPKERLAYEFKGHSLLQAAREADVTRIKKHSLEWVNFKH 360
Qy 435 POTHTETALHCAASAPYPRKQICELLRKGANINEKTEFLPLHVASEKAHNDVVEV 494
Db 361 POTHTETALHCAASAPYPRKQICELLRKGANINEKTEFLPLHVASEKAHNDVVEV 420
Qy 495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEVQOL 554
Db 421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALQMGNEVQOL 480
```



```
QY 555 LQEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSYNCRDIEGRQSTPLHFAAGYNRVS 614
|
|
|
Db 481 LQEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSYNCRDIEGRQSTPLHFAAGYNRVS 540
|
|
|
QY 615 VVEYLQHGADVIAKDKGGLVPLHNACSYGHEVAEALLVKGAVVNVADLWKFTPLHEAA 674
|
|
|
Db 541 VVEYLQHGADVIAKDKGGLVPLHNACSYGHEVAEALLVKGAVVNVADLWKFTPLHEAA 600
|
|
|
QY 675 AKKYEICKLLOHGADPTKKRDGNTPLDLVKDGTDIODLLRGDAALLDAAKKGCGLAR 734
|
|
|
Db 601 AKKYEICKLLOHGADPTKKRDGNTPLDLVKDGTDIODLLRGDAALLDAAKKGCGLAR 660
|
|
|
QY 735 VKLSSPDNVCNCRDQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 794
|
|
|
Db 661 VKLSSPDNVCNCRDQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 720
|
|
|
QY 795 SYGHVDVAALLIKYNACVNATDKWAFPLHEAAKGRTOQLCALLLAHGADPTLKNQEGOT 854
|
|
|
Db 721 SYGHVDVAALLIKYNACVNATDKWAFPLHEAAKGRTOQLCALLLAHGADPTLKNQEGOT 780
|
|
|
QY 855 PLDLVSADYSALLTAAMPSPALPSCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914
|
|
|
Db 781 PLDLVSADYSALLTAAMPSPALPSCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840
|
|
|
QY 915 LDNLGSFSLSVSSSGTREGASSLEKKEVPVGVDFSIOTFVRNLGLEHLMDFPEREQIT 974
|
|
|
Db 841 LDNLGSFSLSVSSSGTREGASSLEKKEVPVGVDFSIOTFVRNLGLEHLMDFPEREQIT 900
|
|
|
QY 975 LDVLVEMGHKELKEIGTINAYCHRRHKLKGVRLISGQOGLNPLYTLNTSGSGTILIDLSP 1034
|
|
|
Db 901 LDVLVEMGHKELKEIGTINAYCHRRHKLKGVRLISGQOGLNPLYTLNTSGSGTILIDLSP 960
|
|
|
QY 1035 DKKEFQSVSEEMOSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEEN 1094
|
|
|
Db 961 DKKEFQSVSEEMOSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEEN 1020
|
|
|
QY 1095 HNHANERMLFHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENSSKNQYVYGIGGTG 1154
|
|
|
Db 1021 HNHANERMLFHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENSSKNQYVYGIGGTG 1080
|
|
|
QY 1155 CPVHKRSCYICHROLFLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGLALAEYV 1214
|
|
|
Db 1081 CPVHKRSCYICHROLFLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNGLALAEYV 1140
|
|
|
QY 1215 IYRGEQAYPEYLITYQIMRPEGWVDG 1240
|
|
|
Db 1141 IYRGEQAYPEYLITYQIMRPEGWVDG 1166
|
|
|
RESULT 7
AAB27211
ID AAB27211 standard; Protein; 1166 AA.
XX
AC AAB27211;
AC
XX 27-FEB-2001 (first entry)
XX
DE Human tankyrase II protein sequence SEQ ID NO: 6.
XX
KW Human; tankyrase II; telomere length; signal transduction.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 124 /note= "encoded by TTA"
FT Misc-difference 125 /note= "encoded by TAC"
XX
XX
PN W0200061813-A1.
XX
PD 19-OCT-2000.
```

```
XX
PF 10-APR-2000; 2000WO-US09558.
XX
PR 09-APR-1999; 99US-0128577.
PR 13-APR-1999; 99US-0129123.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
WPI: 2000-679503/66.
DR N-PSDB; AAC66825.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide
PT encoding the polypeptide useful for modulating or maintaining telomere
PT length, replicative capacity, apoptosis, chromosome packing or gene
PT expression.
XX
XX Claim 4; Fig 4; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase II protein
CC sequence. The protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell.
XX
XX Sequence 1166 AA;
XX
Query Match 94.2%; Score 6092; DB 21; Length 1166;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1164; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 75 MSGRRCCAGGGAACASASAAAEVPAARELFEACRNGDVERVKRLVTPKVNSRDTAGRKST 134
|
|
|
Db 1 MSGRRCCAGGGAACASASAAAEVPAARELFEACRNGDVERVKRLVTPKVNSRDTAGRKST 60
|
|
|
QY 135 PLHFAAGGFRKDVVEYLLQNGANVQARDGGLIPLHNACSGHAEVNVLLRHGADPNAR 194
|
|
|
Db 61 PLHFAAGGFRKDVVEYLLQNGANVQARDGGLIPLHNACSGHAEVNVLLRHGADPNAR 120
|
|
|
QY 195 DNWNTYPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALTDLADPSAKAVLTGEYKKDE 254
|
|
|
Db 121 DNWNTYPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRALTDLADPSAKAVLTGEYKKDE 180
|
|
|
QY 255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 314
|
|
|
Db 181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 240
|
|
|
QY 315 DKGDLVPLHNACSYGHEVETELLVKHGACVNAMDLWQFTPLHEAASKNRVVCSSLISY 374
|
|
|
Db 241 DKGDLVPLHNACSYGHEVETELLVKHGACVNAMDLWQFTPLHEAASKNRVVCSSLISY 300
|
|
|
QY 375 ADPTLLNCHNKSAIDLAPTPOLKERLAYEFKGHSLQAAAREADVTIRKKHLSLEMVNFKH 434
|
|
|
Db 301 ADPTLLNCHNKSAIDLAPTPOLKERLAYEFKGHSLQAAAREADVTIRKKHLSLEMVNFKH 360
|
|
|
QY 435 PQTHETALHCAASPYPKRKOICELELLLRKGANINEKTEFTPLHVASEKAHNDVVEVV 494
|
|
|
Db 361 PQTHETALHCAASPYPKRKOICELELLLRKGANINEKTEFTPLHVASEKAHNDVVEVV 420
|
|
|
QY 495 KHEAKVNALDNLGOTSLHRAAYCGHLOTCRLLSYGCDDPNIIISLOGFTALQNGENVOOL 554
|
|
|
Db 421 KHEAKVNALDNLGOTSLHRAAYCGHLOTCRLLSYGCDDPNIIISLOGFTALQNGENVOOL 480
|
|
|
QY 555 LQEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSYNCRDIEGRQSTPLHFAAGYNRVS 614
|
|
|
Db 481 LQEGISLGNSEADROLLEAAKAGDVETVKKLTCTVQSYNCRDIEGRQSTPLHFAAGYNRVS 540
|
|
|
QY 615 VVEYLQHGADVIAKDKGGLVPLHNACSYGHEVAEALLVKGAVVNVADLWKFTPLHEAA 674
|
|
|
```

Db 541 VVEYLLQHGADVHAKDGGGLVPLHNAACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600  
QY 675 AKGKYEICKLLQHGADPTKKNRDNNTPLDLVKDGDITDIDLLRGDAALLDAAKKGLAR 734  
Db 601 AKGKYEICKLLQHGADPTKKNRDNNTPLDLVKDGDITDIDLLRGDAALLDAAKKGLAR 660  
QY 735 VKLSSPDVNCNCRDTCGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 794  
Db 661 VKLSSPDVNCNCRDTCGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 720  
QY 795 SYGHVDVAALLIKYNACVNAATKWAETPLHEAAQKQRTQCALLLAHGADPTLKNQEGQT 854  
Db 721 SYGHVDVAALLIKYNACVNAATKWAETPLHEAAQKQRTQCALLLAHGADPTLKNQEGQT 780  
QY 855 PLDLVADDDVSALLTAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914  
Db 781 PLDLVADDDVSALLTAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840  
QY 915 LDNLGSFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRLNGLHLMDFEREQIT 974  
Db 841 LDNLGSFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRLNGLHLMDFEREQIT 900  
QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKGVRLISGQGLNPLYTLTNTSGSGTILDLSP 1034  
Db 901 LDVLVEMGHKELKEIGINAYGHRHKLKGVRLISGQGLNPLYTLTNTSGSGTILDLSP 960  
QY 1035 DKEFQSVSEEMOSTVREHRDGHAGGIENRYNLIKQVCNKKLWERTYTHRKEVSEEN 1094  
Db 961 DKEFQSVSEEMOSTVREHRDGHAGGIENRYNLIKQVCNKKLWERTYTHRKEVSEEN 1020  
QY 1095 HNANERMLFHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENSCKSNQYVYGIGGTG 1154  
Db 1021 HNANERMLFHGSPFVNAIHKGFDERHAYIGGMFGAGIYFAENSCKSNQYVYGIGGTG 1080  
QY 1155 CPVHKDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNLALAEYV 1214  
Db 1081 CPVHKDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHHSVTGRPSVNLALAEYV 1140  
QY 1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240  
Db 1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166

RESULT 8

AAY72589

ID AAY72589 standard; Protein: 1166 AA.

AC AAY72589;

XX AAY72589;

XX 02-MAY-2001 (first entry)

XX Human tankyrase homolog protein (THP).

DE Human: tankyrase homolog protein; THP: gene therapy; cancer;

KW tumour; basal cell carcinoma; therapy; genetic mapping;

KW cytostatic.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

FT Misc-difference 368 /label= Met, Leu, Val

FT Misc-difference 392 /label= Asn, Thr

FT Misc-difference 415 /label= Val, Ile

FT Misc-difference 558 /label= Gly, Glu

FT Misc-difference 559 /label= Gly, Ala

FT Misc-difference 764 /label= Leu, Val

FT Misc-difference 884

FT XX /label= Asn, His, Asp, Tyr  
PN WO200104326-A1.  
XX 18-JAN-2001.  
XX 03-JUL-2000; 2000WO-EP06609.  
XX 09-JUL-1999; 99US-0350982.  
XX (PHAA ) PHARMACIA & UPJOHN SPA.  
XX Berthelsen J, Toma S, Isacchi A;  
XX WPI: 2001-168422/17.  
XX N-PSDB; AAD02578.  
XX New tankyrase homolog protein (THP) polynucleotide and polypeptide  
XX useful in gene therapy, diagnosis and treatment or prevention of  
XX unregulated cell growth, such as cancer or tumor cell growth -  
XX Claim 20; Page 50-54; 60pp; English.  
XX The present sequence is human tankyrase homolog  
XX protein (THP). The THP polypeptides and polynucleotides of the  
XX invention are useful in gene therapy and for treating or preventing  
XX unregulated cell growth such as cancer or tumor (e.g. basal cell  
XX carcinoma). The nucleic acid molecules of the invention and their  
XX fragments are useful for restriction fragment length polymorphism  
XX (RFLP) associated with certain disorders, as well as for genetic  
XX mapping. Antisense oligonucleotides, or fragments of nucleic acid  
XX encoding THP are useful as diagnostic tools for probing the  
XX expression of Thp gene in various tissues. THP can be used as  
XX antigens for raising antibodies against them and in assays for  
XX identifying compounds that modulate their activity. They are used in  
XX the manufacture of a medicament directed towards cancers or tumours.  
XX THP are also useful for screening compounds in a variety of drugs  
XX screening techniques and as a research tool for identification,  
XX characterisation and purification of interacting, regulatory proteins.  
XX SQ Sequence 1166 AA;

Query Match

Best Local Similarity 93.6%; Score 6053; DB 22; Length 1166;

Mismatches 1157; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 75 MSRRRCAGGAGACASAAAEPAARELFEACRNGDVERVKRLVTEPKVNSRDTAGRKST 134

Db 1 MSRRRCAGGAGACASAAAEPAARELFEACRNGDVERVKRLVTEPKVNSRDTAGRKST 60

QY 135 PLHFAAGFGKRDVVEYLLONGANVOARDGGGLPLHNAACSYGHYEVAELLVKHGADPNAR 194

Db 61 PLHFAAGFGKRDVVEYLLONGANVOARDGGGLPLHNAACSYGHYEVAELLVKHGADPNAR 120

QY 195 DNWNVTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 254

Db 121 DNWNVTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKKDE 180

QY 255 LLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHIAAGYNRVKIVQLLQHGADVHAK 314

Db 181 LLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHIAAGYNRVKIVQLLQHGADVHAK 240

QY 315 DKGDLVPLHNAACSYGHYEVTLLVKGACVNAAMDQWFTPLHEAASKNRVEVCSLLLSYG 374

Db 241 DKGDLVPLHNAACSYGHYEVTLLVKGACVNAAMDQWFTPLHEAASKNRVEVCSLLLSYG 300

QY 375 ADPTLLNCHNKSAIDLAPTLPOLKERLAYEFKSHLSLQAAAREADVTRIKKHSLEMYNFKH 434

Db 301 ADPTLLNCHNKSAIDLAPTLPOLKERLAYEFKSHLSLQAAAREADVTRIKKHSLEMYNFKH 360

QY 435 POTHETALHCAASAPYPRKQICELLRLRGANINEXTKFELTPLHVASEKAHNDVVEVVV 494

Db 361 POTHETALHCAASAPYPRKQICELLRLRGANINEXTKFELTPLHVASEKAHNDVVEVVV 420

```
QY 495 KHEAKVNALDNLGOTSLHRAAYCGHLCOTCRLLLSYGCDPNIIISLOGFTALQMGNEVVOOL 554
DB 421 KHEAKVNALDNLGOTSLHRAAYCGHLCOTCRLLLSYGCDPNIIISLOGFTALQMGNEVVOOL 480
QY 555 LOEGISLGNSEADROLLLEAAKAGDVETVKKICTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
DB 481 LOEGISLGNSEADROLLLEAAKAGDVETVKKICTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540
QY 615 VVEYLLOHGADYHAKDKGLVPLHNACSYGHVEAEALLVKHGVNVADLWKFTPLHFAA 674
DB 541 VVEYLLOHGADYHAKDKXXLVPPLHNACSYGHVEAEALLVKHGVNVADLWKFTPLHFAA 600
QY 675 AKGKEICKLLQLHGADPTKKNRDNTPDLVLDKDGDTDIQDLRLRGDAALLDAAKKGCCLAR 734
DB 601 AKGKEICKLLQLHGADPTKKNRDNTPDLVLDKDGDTDIQDLRLRGDAALLDAAKKGCCLAR 660
QY 735 VKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHNA 794
DB 661 VKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHNA 720
QY 795 SYGHVDVAALLIKYNACVNATDKWAFTPLHAAQKGRQOLCALLAHAGADPTLKNQEGQT 854
DB 721 SYGHVDVAALLIKYNACVNATDKWAFTPLHAAQKGRQOLCALLAHAGADPTLKNQEGQT 780
QY 855 PLDLVSADDSVALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914
DB 781 PLDLVSADDSVALLTAAMPSPALPCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840
QY 915 LNLSSFSFELSIVSSSTGEGASLEKKEVPCVDFSTQFVNRNGLGHEHMDIFEREQIT 974
DB 841 LNLSSFSFELSIVSSSTGEGASLEKKEVPCVDFSTQFVNRNGLGHEHMDIFEREQIT 900
QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPLYTLNTSGSGTILIDLSP 1034
DB 901 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPLYTLNTSGSGTILIDLSP 960
QY 1035 DKEFQSVSEEMQSVREHRDGGHAGGIFNRYNLIKIOKVCNKKLWERYTHRRKRVSEEN 1094
DB 961 DKEFQSVSEEMQSVREHRDGGHAGGIFNRYNLIKIOKVCNKKLWERYTHRRKRVSEEN 1020
QY 1095 HNHNANERMLFHGSPFVNAIIHKGFDERHAYIGGMFAGIYFAENSCKSNQYVYIGGGTG 1154
DB 1021 HNHNANERMLFHGSPFVNAIIHKGFDERHAYIGGMFAGIYFAENSCKSNQYVYIGGGTG 1080
QY 1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
DB 1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140
QY 1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
DB 1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166

RESULT 9
AAY97748
ID AAY97748 standard; Protein; 1100 AA.
XX
AC AAY97748;
XX
DT 06-AUG-2001 (first entry)
DE
DE Tankyrase homologue isotype 1 protein sequence.
XX
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
```

```
XX
PF 25-OCT-2000; 2000WO-US41528.
XX
PR 25-OCT-1999; 99US-0427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
XX WPI: 2001-300503/31.
XX N-PSDB; AAA91487.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for
PT inducing or preventing cell proliferation in cells, and for diagnosing,
PT treating or preventing cell cycle associated disorders such as cancer
PT
XX
PS Claim 22; Fig 3: 63pp; English.
XX
XX This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein
XX of the invention. The invention also relates to the TaHo-2 protein.
XX The TaHo proteins are useful for inducing or preventing cell
XX proliferation in cells, and in the study or treatment of conditions
XX mediated by the cell cycle proteins, such as to diagnose, treat or
XX prevent cell cycle associated disorders, preferably cancer. The TaHo
XX coding sequences are useful as hybridisation probes, in chromosome and
XX gene mapping and in the generation of anti-sense DNA and RNA. The coding
XX sequences are also useful for the preparation of TaHo, for generating
XX either transgenic animals or knock out animals which, in turn, are useful
XX in a development and screening of therapeutically useful agents, in gene
XX therapy, as vaccine, and for construction of hybridisation probes for
XX mapping the gene which encodes TaHo and for the genetic analysis of
XX individuals with genetic disorders. The TaHo proteins, and their coding
XX sequences are useful in screening assays.
XX
SQ Sequence 1100 AA:
Query Match 89.2%; Score 5766; DB 22; Length 1100;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 141 GFGKRDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYIT 200
DB 1 GFGKRDVVEYLLQNGASVQARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNYIT 60
QY 201 PLHEAAIKGIDVCILVLOHGAETIRNTDGTALDLDADPSAKAVLTGEYKDELLESAR 260
DB 61 PLHEAAIKGIDVCILVLOHGAETIRNTDGTALDLDADPSAKAVLTGEYKDELLESAR 120
QY 261 SGNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADYHAKDKGDLV 320
DB 121 SGNEEKMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADYHAKDKGDLV 180
QY 321 PLHNACSYGHVEYVELLYVKHGCYNAMDLPQFTPLHFAASKNRVEVCSLLSYGADPTLL 380
DB 181 PLHNACSYGHVEYVELLYVKHGCYNAMDLPQFTPLHFAASKNRVEVCSLLSYGADPTLL 240
QY 381 NCHNKSADLAPTQPKERLAYEFKSHLSLQAAAREADVTRIKKHLSEMVNFKHPQTHET 440
DB 241 NCHNKSADLAPTQPKERLAYEFKSHLSLQAAAREADVTRIKKHLSEMVNFKHPQTHET 300
QY 441 ALHCAAAAPYKPKKOICELLRLKGANINEKTEFTPLHVAASEKHAHNDVVEVVKHEAKV 500
DB 301 ALHCAAAAPYKPKKOICELLRLKGANINEKTEFTPLHVAASEKHAHNDVVEVVKHEAKV 360
QY 501 NALDNLGOTSLHRAAYCGHLCOTCRLLLSYGCDPNIIISLOGFTALQMGNEVVOOLQEGIS 560
DB 361 NALDNLGOTSLHRAAYCGHLCOTCRLLLSYGCDPNIIISLOGFTALQMGNEVVOOLQEGIS 420
QY 561 LGNSEADROLLLEAAKAGDVETVKKICTVQSVNCRDIEGRQSTPLHFAAGYNRVSWEYLL 620
DB 421 LGNSEADROLLLEAAKAGDVETVKKICTVQSVNCRDIEGRQSTPLHFAAGYNRVSWEYLL 480
```

QY 621 QHGADVHAKDGGGLVPLHNACSYGHEVEAELLVKHGAVVNVADLWKFTPLHEAAAKGYE 680  
 DB 481 QHGADVHAKDGGGLVPLHNACSYGHEVEAELLVKHGAVVNVADLWKFTPLHEAAAKGYE 540  
 QY 681 ICKLLQHGADPTKKNRDNTPDLVDKGGTDIDLLRGDAALLDAAKKGCLARVKKLSS 740  
 DB 541 ICKLLQHGADPTKKNRDNTPDLVDKGGTDIDLLRGDAALLDAAKKGCLARVKKLSS 600  
 QY 741 PDVNCRDGTGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800  
 DB 601 PDVNCRDGTGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660  
 QY 801 VAALLIKYNACVNATDKWFTPLHEAAQKGTOLCALLLAHAGADPTLKNGEGOTPLDLVS 860  
 DB 661 VAALLIKYNACVNATDKWFTPLHEAAQKGTOLCALLLAHAGADPTLKNGEGOTPLDLVS 720  
 QY 861 ADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSSPSLSAASLDNLG 920  
 DB 721 ADDVSALLTAAMPSPALPCYKQVNLGVRSFGATADALSSGSSPSLSAASLDNLG 780  
 QY 921 SFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVE 980  
 DB 781 SFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLGLEHLMDFEREQITLDVLVE 840  
 QY 981 MGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPYLTNTSGSGTILIDLPDDKEFQ 1040  
 DB 841 MGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPYLTNTSGSGTILIDLPDDKEFQ 900  
 QY 1041 SVEEEMOSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERTHRKVESEENHNHANE 1100  
 DB 901 SVEEEMOSTVREHRDGGHAGGIFNRYNLIKIQVCNKKLWERTHRKVESEENHNHANE 960  
 QY 1101 RMLFHSGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPVHKD 1160  
 DB 961 RMLFHSGSPFVNAIHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPVHKD 1020  
 QY 1161 RSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYIRGEQ 1220  
 DB 1021 RSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPGHHSVTGRPSVNGLALAEYVYIRGEQ 1080  
 QY 1221 AYPEYLITYQIMRPEGMVDG 1240  
 DB 1081 AYPEYLITYQIMRPEGMVDG 1100

RESULT 10  
 AAY05734

ID AAY05734 standard; Protein; 1074 AA.

AC AAY05734;

XX 19-JUL-1999 (first entry)

DE Human Grb7 effector 2.2412 protein.

KW Grb7 effector; 2.2412 protein; human; signal transduction;  
 KW tumour marker; breast cancer; prostate cancer; prognosis;  
 KW diagnosis.

XX Homo sapiens.

XX W09915647-A1.

XX 01-APR-1999.

XX 23-SEP-1998; 98WO-AU00795.

XX 23-SEP-1997; 97AU-0009388.

XX (GARV-) GARVAN INST MEDICAL RES.

XX Daly RJ, Sutherland RL;

XX

DR WPI; 1999-254707/21.  
 XX N-PSDB; AAX25366.

PT New candidate effector for the Grb7 family of signalling proteins,  
 PT and specific antibody, useful for detection and treatment of cancer  
 XX Claim 9; Fig 1; 24pp; English.

XX The present sequence represents a novel candidate effector for  
 CC the Grb7 family of signalling proteins, termed 2.2412. The  
 CC sequence is predicted from a partial cDNA (see AAX25366). 2.2412  
 CC mRNA transcripts (7 kb) were detected in all tissues examined  
 CC with the exception of kidney. Expression was particularly high in  
 CC skeletal muscle and placenta. Analysis of the sequence revealed  
 CC significant homology to a large number of proteins containing  
 CC ankyrin-like repeats. The 2.2412 gene was localised to between  
 CC chromosome 10q33.2 and proximal 10q33.32. Deletions in the  
 CC 10q22-25 region have been detected in human breast, prostate,  
 CC renal, small cell lung and endometrial carcinomas, glioblastoma  
 CC multiforme, melanoma and meningiomas. Detection of the protein  
 CC encoded by the 2.2412 cDNA in a sample should provide a useful  
 CC tumour marker and/or prognostic indicator for certain human  
 CC cancers, in particular breast cancer and prostate cancer.  
 CC Antagonism of the interaction between Grb7 family members and the  
 CC encoded protein should provide a novel treatment strategy for human  
 CC diseases exhibiting aberrant receptor tyrosine kinase signalling,  
 CC such as cancer. Anti-Grb7 antibodies can be used in methods of  
 CC detecting the presence of 2.2412 protein in a sample.

XX Sequence 1074 AA:

Query Match 86.7%; Score 5605; DB 20; Length 1074;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1069; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 167 IPLHNACSFSGHAEVNNLLLRHGADPNARDNNWNTPLHEAAIKGKIDVICIVLLQHGAEPTI 226  
 DB 1 IPLHNACSFSGHAEVNNLLLRHGADPNARDNNWNTPLHEAAIKGKIDVICIVLLQHGAEPTI 60  
 QY 227 RNTDGTALDADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 286  
 DB 61 RNTDGTALDADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 120  
 QY 287 TPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVPLHNACSYGHEVEYTELLVXHGACVNA 346  
 DB 121 TPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVPLHNACSYGHEVEYTELLVXHGACVNA 180  
 QY 347 MDLWQFTPLHEAASKNRVEVCULLSYGADPTLLNCHNKSAIDLAPTPOLKERLAYEFKG 406  
 DB 181 MDLWQFTPLHEAASKNRVEVCULLSYGADPTLLNCHNKSAIDLAPTPOLKERLAYEFKG 240  
 QY 407 HSLLOAAREADVTRIKKHLSELMVNFKHPOTHETALHCAASVPKPKQICELLLRKGAN 466  
 DB 241 HSLLOAAREADVTRIKKHLSELMVNFKHPOTHETALHCAASVPKPKQICELLLRKGAN 300  
 QY 467 INEKTKEFTPLHVASEKAHNDVVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL 526  
 DB 301 INEKTKEFTPLHVASEKAHNDVVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL 360  
 QY 527 LSYGCDPNIIISLOGFTALQMGNEVVOQLQEGISLGNSEADROLLEAAKAGDVETVKLC 586  
 DB 361 LSYGCDPNIIISLOGFTALQMGNEVVOQLQEGISLGNSEADROLLEAAKAGDVETVKLC 420  
 QY 587 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 646  
 DB 421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480  
 QY 647 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPLDLV 706  
 DB 481 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGYEICKLLQHGADPTKKNRDNTPLDLV 540  
 QY 707 KUGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDVNCRDPTQGRHSTPLHLAAGYNNL 766  
 DB 540 KUGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDVNCRDPTQGRHSTPLHLAAGYNNL 766



||||| 384 LRKGANINEKTKPYPEFLTPHASEKANDVVEVVVHKAVALNDNLGOTSLHRAAYC 443  
||||| 518 GHLOTCRLLLSYGCDPNIIISLOGFTALOMGNENVOQLL-----QEGI 559  
||||| 444 GHLOTCRLLLSYGCDPNIIISLOGFTALOMGNENVOQLLTVKKLCIVQSVNCRDIEGQEI 503  
QY 560 SLGNSEADROLLEAAKAGDVEVVKLCIVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619  
||||| 504 SLGNSEADROLLEAAKAGDVE-----RQSTPLHFAAGYNRVSVVEYL 545  
QY 620 LOHGADVHAKDGLVPLHNAACSYGHEVAELLVKGAVVNVADLWKFTPLHEAAKAGKY 679  
||||| 546 LOHGADVHAKDGLVPLHNAACSYGHEVAELLVKGAVVNVADLWKFTPLHEAAKAGKY 605  
QY 680 EICKLLLOHGADPTKKNRDGNTPDLVDKGDIDIDLLRGAALLDAKKGCLARVKKLS 739  
||||| 606 EICKLLLOHGADPTKKNRDGNTPDLVDKGDIDIQ-----RVKKLS 646  
QY 740 SPDVNCRTDQ-----RHSTPLHLAAGYNLEVAEYLLQHGADV 780  
||||| 647 SPDVNCRTDQDGLLRGDAALLDAKKGCLARHSTPLHLAAGYNLEVAEYLLQHGADV 706  
QY 781 AODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAKGRTOQLCALLA 840  
||||| 707 AODKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAKGRTOQLCALLA 766  
QY 841 HGADPTLKNQEGOTPLDVSADVDVSAALLTAAMPSPALPCYKPOVLNGVRSPGATADALS 900  
||||| 767 HGADPTLKNQEGOTPLDVSADVDVSAALLTAAMPSPALPCYKPOVLNGVRSPGATADALS 826  
QY 901 SGPSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLG 960  
||||| 827 SGPSPSSLSAASSLDNLGSGFSELSSVSSSGTEGASSLEKKEVPGVDFSIQFVRNLG 886  
QY 961 LEHLMDFIREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTL 1020  
||||| 887 LEHLMDFIREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTL 946  
QY 1021 NTSGSGTILDLSPDQKEFQSVSEEMOSTVREHRRGGHAGGIFNRYNLIKIOKVCNKKLW 1080  
||||| 947 NTSGSGTILDLSPDQKEFQSVSEEMOSTVREHRRGGHAGGIFNRYNLIKIOKVCNKKLW 1006  
QY 1081 ERYTHRRKEVSEENHNHANERMLFHGSPFNVAIIHKGFDERHAYIGMGFGAGIYFAENSS 1140  
Db 1007 ERYTHRRKEVSEENHNHANERMLFHGSPFNVAIIHKGFDERHAYIGMGFGAGIYFAENSS 1066  
QY 1141 KSNQYVYIGGGTGPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVT 1200  
Db 1067 KSNQYVYIGGGTGPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPPGHHSVT 1126  
QY 1201 GRPSVNGLALAEYVYRGOAYPEYLITYQIMRPEGWVG 1240  
Db 1127 GRPSVNGLALAEYVYRGOAYPEYLITYQIMRPEGWVG 1166

RESULT 12

ID AAB27210  
XX AAB27210 standard; Protein; 1431 AA.

AC AAB27210;

DT 27-FEB-2001 (first entry)

XX Human tankyrase II protein sequence SEQ ID NO: 4.

DE Human; tankyrase II; telomere length; signal transduction.

XX Homo sapiens.

FW Key Location/Qualifiers

FT Misc-difference 41 /label= Xaa

FT Misc-difference 97 /note= "encoded by ANG"  
FT /label= Xaa  
FT /note= "encoded by ANG"  
FT Misc-difference 109 /label= Xaa  
FT /note= "encoded by NCG"  
FT Misc-difference 113 /label= Xaa  
FT /note= "encoded by ANG"  
FT /label= Xaa  
FT /note= "ANK domain"  
FT Domain 122..935  
FT Misc-difference 125 /label= Xaa  
FT /note= "encoded by AAN"  
FT Misc-difference 138 /label= Xaa  
FT /note= "encoded by NTC"  
FT Misc-difference 165 /label= Xaa  
FT /note= "encoded by NGT"  
FT Misc-difference 172..173 /label= XaaXaa  
FT /note= "encoded by NTTRAA"  
FT Misc-difference 177 /label= Xaa  
FT /note= "encoded by CAN"  
FT Misc-difference 184 /label= Xaa  
FT /note= "encoded by NCA"  
FT Misc-difference 187 /label= Xaa  
FT /note= "encoded by NAT"  
FT Misc-difference 201 /label= Xaa  
FT /note= "encoded by CMT"  
FT Misc-difference 203..204 /label= XaaXaa  
FT /note= "encoded by RAAANT"  
FT Misc-difference 206 /label= Xaa  
FT /note= "encoded by NAT"  
FT Misc-difference 210 /label= Xaa  
FT /note= "encoded by CNA"  
FT Misc-difference 212 /label= Xaa  
FT /note= "encoded by NGT"  
FT Misc-difference 214 /label= Xaa  
FT /note= "encoded by NAM"  
FT Misc-difference 226..227 /label= XaaXaa  
FT /note= "encoded by CNCNAT"  
FT Misc-difference 236..237 /label= XaaXaa  
FT /note= "encoded by GANNNT"  
FT Misc-difference 666 /label= Xaa  
FT /note= "encoded by TNT"  
FT Misc-difference 724 /label= Xaa  
FT /note= "encoded by AAN"  
FT Misc-difference 726 /label= Xaa  
FT /note= "encoded by CNT"  
FT Misc-difference 734 /label= Xaa  
FT /note= "encoded by GAN"  
FT Misc-difference 740 /label= Xaa  
FT /note= "encoded by NTC"  
FT Misc-difference 762..763

FT /label= XaaXaa  
 FT /note= "encoded by TTNTNT"  
 FT 947..993  
 FT Domain  
 FT /note= "SAM domain"  
 FT Misc-difference 992  
 FT /label= Xaa  
 FT /note= "encoded by NAG"  
 FT Misc-difference 1267  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1274  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1277  
 FT /label= Xaa  
 FT /note= "encoded by TGA"  
 FT Misc-difference 1279  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1294  
 FT /label= Xaa  
 FT /note= "encoded by TGA"  
 FT Misc-difference 1314  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1322..1323  
 FT /label= XaaXaa  
 FT /note= "encoded by TAATAA"  
 FT Misc-difference 1353  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1358  
 FT /label= Xaa  
 FT /note= "encoded by TGA"  
 FT Misc-difference 1371  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1377  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1380  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1393  
 FT /label= Xaa  
 FT /note= "encoded by TCA"  
 FT Misc-difference 1399  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1403  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT Misc-difference 1408  
 FT /label= Xaa  
 FT /note= "encoded by TAG"  
 FT Misc-difference 1419  
 FT /label= Xaa  
 FT /note= "encoded by TAA"  
 FT /note= "Xaa-unknown"  
 XX WO200061813-A1.  
 XX  
 XX 19-OCT-2000.  
 PD  
 XX 10-APR-2000; 2000WO-US09558.  
 PF  
 XX 09-APR-1999; 9905-0128577.  
 PR  
 XX 13-APR-1999; 9905-0129123.  
 XX  
 XX (GERO-) GERON CORP.  
 PA  
 XX Morin GB, Funk WD, Piatyszek MA;  
 PI  
 XX

DR WPI; 2000-679503/66.  
 DR N-PSDB; AAC66824.  
 XX  
 PT Novel mammalian Tankyrase II polypeptide and the polynucleotide  
 FT encoding the polypeptide useful for modulating or maintaining telomere  
 PT length, replicative capacity, apoptosis, chromosome packing or gene  
 PT expression -  
 XX  
 PS Disclosure; Fig 3; 52pp; English.  
 XX  
 CC The present sequence is a version of the human tankyrase II protein  
 CC sequence. The protein is thought to be involved in signal transduction in  
 CC the cell, and to have binding activity for other telomere-associated  
 CC proteins. It is possible that it plays a role in the regulation of  
 CC telomere length, thus affecting the replicative ability of the cell. The  
 CC protein is useful for ribosylating target proteins, for determining  
 CC tankyrase II binding activity in a sample, and for modulating telomere  
 CC length in a cell.  
 XX  
 SQ Sequence 1431 AA;  
 Query Match 84.7%; Score 5474; DB 21; Length 1431;  
 Best Local Similarity 86.6%; Pred. No. 0;  
 Matches 1076; Conservative 6; Mismatches 157; Indels 4; Gaps 2;  
 QY 1 RCSARRGAAGGQAGQARGARVGAAGHGTAPDPTAGSQAARALSASSPGGLALLAGPGLL 60  
 DB 25 RCLRRGAAGGQAGHXRGCARGHGTAPDPTAGSQAARALSASSPGGLALLAGPGLL 84  
 QY 61 RLLALLAVAAARIMSGRRCAAGGGAACASAAAEVPAARELFACRNGDVERVKRLVTP 120  
 DB 85 RLLALLAVAAAXIMSGRRCAAGGGAACASAAAEVPAARELFACRNGDVERKKLVTP 144  
 QY 121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLONGANVQARDGGLIPLHNACSFGEAEV 180  
 DB 145 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLONGANVQARDGGLIPLHNACSFGEAXX 204  
 QY 181 VNLLRHGADPNARDNNWYTPPLHAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLADP 240  
 DB 205 IXLLXHXAXPNARDNNWYTPXEEAIIKGIKXCVLLQHGAEPTIRNTDGTALDLADP 264  
 QY 241 SAKAVLTGEYKKDELLESARSNEEKMALITPLNVNCHASDGRKSTPLHAAAGVNRVKI 300  
 DB 265 SAKAVLTGEYKKDELLESARSNEEKMALITPLNVNCHASDGRKSTPLHAAAGVNRVKI 324  
 QY 301 VOLLLQHGADYHAKDKDGLVPLHNACSYCHYEVTETLLVKGACVNMADLWQFTPLHEAAS 360  
 DB 325 VOLLLQHGADYHAKDKDGLVPLHNACSYCHYEVTETLLVKGACVNMADLWQFTPLHEAAS 384  
 QY 361 KNRVEVCSLLISYGADPTLLNCHNKSADLAPTOLKERLAYEFKGHSLLOAAREADVTR 420  
 DB 385 KNRVEVCSLLISYGADPTLLNCHNKSADLAPTOLKERLAYEFKGHSLLOAAREADVTR 444  
 QY 421 IKKHLISLEMVNFKHPQTHETALHCAAAASPYPKRKOICELLURKGANINEKTEFTPLHV 480  
 DB 445 IKKHLISLEMVNFKHPQTHETALHCAAAASPYPKRKOICELLURKGANINEKTEFTPLHV 504  
 QY 481 ASEKAHNDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHLQTCRLLLSYGCDPNTIISLQ 540  
 DB 505 ASEKAHNDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHLQTCRLLLSYGCDPNTIISLQ 564  
 QY 541 FTALQMGNEVVOOLLQEGISLGNSEADROLLEAAKAGDVETVKLCTVQSVNCRDIEGRQ 600  
 DB 565 FTALQMGNEVVOOLLQEGISLGNSEADROLLEAAKAGDVETVKLCTVQSVNCRDIEGRQ 624  
 QY 601 STPLHFAAGYNRVSVVEYLLQHGADYHAKDKGGLVPLHNACSYGHEVAELLVKGAAYN 660  
 DB 625 STPLHFAAGYNRVSVVEYLLQHGADYHAKDKGGLVPLHNACSYGHEVAELLVKGAAYN 684  
 QY 661 VADLWKFTPLHEAAAKGKYEICKLLOLQHGADPTKKNRDGNTPDLVYKDGDTDIQDLRGD 720  
 DB 685 VADLWKFTPLHEAAAKGKYEICKLLOLQHGADPTKKKKKXIXLDLVKDGXDTDIQDLRGD 744

QY 721 AALLDAARKGCLARYKKLSSPDNVCNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780  
 Db 745 AVLLDAARKGCLARVKKXXEFPDNYNCRDTQGRHSTPLHL-AGXXXXXXXKXXXXXXXXXX 803  
 QY 781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPLHFAAQKGTQLCALLLA 840  
 Db 804 XXX 863  
 QY 841 HGADPTLKNBGOPTDLVLSADDDVALLTAAMPSPALPSCYKPOVLNGVRSPGATADALS 900  
 Db 864 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTAMPSPVLPSCNKQVLNGVRSPGATADALS 923  
 QY 901 SGSPSSLSAASLDNLGSGFSLSVSSGTEGASSLEKKEVPGVDFPSITOFVRLNG 960  
 Db 924 SGSPSSLSAASLDNLGSGFSLSVSSGTEGASSLEKKEVPGVDFPSITOFVRLNG 983  
 QY 961 LEHLMDFERQITDLVLMVGHKELKEIGINAYGHRHKLKIGVERLISGOQGNPYLTL 1020  
 Db 984 LEHLMDFXREQITDLVLMVGHKELKEIGINAYGHRHKLKISFERLISGOQGNPYLTL 1043  
 QY 1021 NTSGSGTILDLSPDDKEFQSVSEEMOSTVREHRDGGHAGGIFNRYNLIKIQVCNK--- 1077  
 Db 1044 NTSGSGTILDLSPDDKEFQSVSEEMOSTVREHRDGGHAGGIFNRYNLIKIQVCNRAKI 1103  
 QY 1078 KIWERYTHRRKEVSEENHNHANERMLFHSPPFYNAIILHKGFDERHAYIGGMFGAGIYFAE 1137  
 Db 1104 RHEERYTHRRKEVSEENHNHANERMLFHSPPFYNAIILHKGFDERHAYIGGMFGAGIYFAE 1163  
 QY 1138 NSSKSNQYVYGIGGTCPCVHKRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGPHH 1197  
 Db 1164 NSSKSNQYVYGIGGTCPCVHKRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGPHH 1223  
 QY 1198 SVTGRPSVNGLALAEYIYVYGEQAYPEYLITYQIMRPEGWVDG 1240  
 Db 1224 SVTGRPSVNGLALAEYIYVYGEQAYPEYLITYQIMRPEGWVDG 1266

## RESULT 13

AAB27212  
 ID AAB27212 standard; Protein; 1327 AA.  
 XX  
 AC AAB27212;  
 DT 27-FEB-2001 (first entry)  
 XX  
 DE Human tankyrase I protein sequence SEQ ID NO: 8.  
 XX  
 KW Human; tankyrase II; telomere length; signal transduction.

XX Homo sapiens.  
 OS Wo200061813-A1.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 10-APR-2000; 2000WO-US09558.  
 XX  
 PR 09-APR-1999; 99US-0128577.  
 PR 13-APR-1999; 99US-0129123.  
 XX  
 PA (GERO-) GERON CORP.  
 XX  
 PI Morin GB, Funk WD, Piatyszek MA;  
 XX WPI; 2000-679503/66.

PT Novel mammalian Tankyrase II polypeptide and the polynucleotide  
 encoding the polypeptide useful for modulating or maintaining telomere  
 length, replicative capacity, apoptosis, chromosome packing or gene  
 expression

PS Claim 4; Fig 5; 52pp; English.

XX

CC The present sequence is a version of the human tankyrase I protein  
 sequence. The invention relates to the isolation of the protein and  
 CC coding sequences of human tankyrase II. This protein is thought to be  
 CC involved in signal transduction in the cell, and to have binding activity  
 CC for other telomere-associated proteins. It is possible that it plays a  
 CC role in the regulation of telomere length, thus affecting the replicative  
 CC ability of the cell. The protein is useful for ribosylating target  
 CC proteins, for determining tankyrase II binding activity in a sample, and  
 CC for modulating telomere length in a cell.

XX Sequence 1327 AA;

Query Match 79.0%; Score 5103.5; DB 21; Length 1327;  
 Best Local Similarity 79.2%; Pred. No. 0;  
 Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAHGTADPVTAGSQAARALASASSPGLLALLAGPGLLLRLLALLAVAAARIMSGRCA 81  
 Db 112 SAAGVAPNPAGSGNNPSSSSPTSSSSSPSPG-----SSLAESPEAGYSSTAPL 165

QY 82 GGGACASAAAAEVEPAARELFACRNGDVERVKRLVTPPEKVNRSRDTAGRKSTPLHFAAG 141  
 Db 166 GPGAAGPCTGVPVSGALRELLEACRNGDYSVRKRLVDAANVNNAKDMGRKSSPLHFAAG 225

QY 142 FGRKDVVEYLLQNGANYOARDGGLIPLHNACSFHAEVNNLLRHGADPNARDNNWYTP 201  
 Db 226 FGRKDVVEYLLQNGANYOARDGGLIPLHNACSFHAEVNNLLRHGADPNARDNNWYTP 285

QY 202 LHEAAIKGKIDVCLVLLQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKDELLESARS 261  
 Db 286 LHEAAIKGKIDVCLVLLQHGADPNIRNTDGTALDLADPSAKAVLTGEYKDELLESARS 345

QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVIVQLLQHGADVHAKDKGLVP 321  
 Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVIVQLLQHGADVHAKDKGLVP 405

QY 322 LHNACSYGHYEVELLKVHGACVNMALWQFTPLHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 381  
 Db 406 LHNACSYGHYEVELLKVHGACVNMALWQFTPLHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 465

QY 382 CHNKSAITDAPTPOLKRLAYEFKSHLSLQAAAREADYTRIKKHLSEWVNFKHPQTHETA 441  
 Db 466 CHKSADMAPTPELRRLTYEFKSHLSLQAAAREADYTRIKKHLSEWVNFKHPQTHETA 525

QY 442 LHCAASPYPRKQICELLKRGANINEKTEFPLTHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 501  
 Db 526 LHCAASPYPRKQICELLKRGANINEKTEFPLTHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 585

QY 502 ALDNLGOTSLHRAAYCCHLQTCRLLLSYGCDDPNIIISLQGTALOMGNVVOQLLOEISL 561  
 Db 586 ALDNLGOTSLHRAAYCCHLQTCRLLLSYGCDDPNIIISLQGTALOMGNVVOQLLOEISL 645

QY 562 GNSEADROLLEAAKAGDVTETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 621  
 Db 646 RTSDVDYRLLEASKAGDLETVKQICSSQNVNCRDIEGRQSTPLHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 705

QY 622 HGADVHAKDKGGLVPLHNACSYGHYEVELLKVHGACVNMALWQFTPLHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 681  
 Db 706 HGADVHAKDKGGLVPLHNACSYGHYEVELLKVHGACVNMALWQFTPLHFAAGYNNRVIVQLLQHGADVHAKDKGLVP 765

QY 682 CKLLQHGADPTKKNRDGNTPLDVKDGTDIQDLRGDAALLDAKKGCLARVKLSSP 741  
 Db 766 CKLLQHGADPTKKNRDGNTPLDVKDGTDIQDLRGDAALLDAKKGCLARVKLSSP 825

QY 742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801  
 Db 826 ENINCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 885

QY 802 AALLIKYNACVNATDKWAFPLHFAAQKGTQLCALLLAHGADPTLKNBGOPTDLVLSA 861  
 Db 886 AALLIKYNACVNATDKWAFPLHFAAQKGTQLCALLLAHGADPTLKNBGOPTDLVLSA 945

QY 862 DVVSALLTAAMPSPALPSCYKPOVLNGVRSPGATADALSPPSSLSAASLDNLGSGS 921





Db 886 AALLIKYNTCVNATDKWAFPLHEAAQGRQLCALLAHGADPTMKNQEGQTPDLATA 945  
 QY 862 DDVSAALLTAAMPSPALPSCYKQVNLNGVRSPGATADALSQSPSSLSAASSLDNLGSG 921  
 Db 946 DDIRALLIDAMPPEALPTCFKPOAT-----VVSASLISPASTPSCLSAASSIDNLGTP 998  
 QY 922 FSELSSVVSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDFIREQITLDLV 979  
 Db 999 LAELAVGASNAGDGAAGTERKECEVAGLDNLSQFLKSLGLEHLRDFITEQITLDVLA 1058  
 QY 980 EMGHKELKEIGINAYGHRHKLIGVERLISGQOGLNPLYLTNTSGSGTILIDLSPDKDEF 1039  
 Db 1059 DMGHEELKEIGINAYGHRHKLIGVERLISGQOGLNPLYLTNTSGSGTILIDLAPDEKEY 1118  
 QY 1040 QSVVEEMQSTVREHRDGHAGGFNRYNLIKQVCKNKLWERYTHRRKVESENNHNN 1099  
 Db 1119 QSVVEEMQSTVREHRDGHAGGFNRYNLIKQVCKNKLWERYTHRRKVESENNHNN 1178  
 QY 1100 ERMFLHGSFPFNALIIHKGDFDERHAYIGMFGAGIYFAENSCKSNQYVYGGGTGCPVHK 1159  
 Db 1179 ERMFLHGSFPFNALIIHKGDFDERHAYIGMFGAGIYFAENSCKSNQYVYGGGTGCPVHK 1238  
 QY 1160 DRSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPPGHSHSVTGRPSVNGLALAEVYIYRGE 1219  
 Db 1239 DRSCYICHRQLLCRVTLGKSFLOFSAMKMAHSPPGHSHSVTGRPSVNGLALAEVYIYRGE 1298  
 QY 1220 QAYPEYLITYQIMRPE 1235  
 Db 1299 QAYPEYLITYQIMRPE 1314

## RESULT 15

AAB66279  
 ID AAB66279 standard; Protein; 1327 AA.

AC AAB66279;  
 XX  
 DT 05-APR-2001 (first entry)  
 DE Human tankyrase1 SEQ ID NO: 4.  
 KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasia; aging;  
 KW Inflammatory disorder.  
 XX Homo sapiens.  
 XX WO200100849-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 XX 28-JUN-2000; 2000WO-0517827.  
 XX  
 XX 29-JUN-1999; 99US-0141582.  
 XX (ICOS-) ICOS CORP.  
 PA  
 PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 XX  
 XX WPI: 2001-102896/11.  
 DR N-PSDB; AAF63838.  
 XX

New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -

Example 1; Page 118-121; 242pp; English.

The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of

CC telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

XX  
 SQ Sequence 1327 AA;  
 Query Match 79.0%; Score 5103.5; DB 22; Length 1327;  
 Best Local Similarity 79.2%; Pred. No. 0;  
 Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAGTAPDPVTVAGTAAARALSSAPGGIALLLAGPGLLLRLLALLLAAVAARIMSGRRCA 81  
 Db 112 SAAGVAPNPAGSGNNPSSSSPSSSTSSSSPSPG-----SSLAESPEAAGVSTAPL 165  
 QY 82 GGGACASAAAEAVEPAARELFEACRNGDVERVKRLVTPKVNRSRDTAGRKSTPLHFAAG 141  
 Db 166 GPGAAGPGTGPVAVSGALRELLEACRNGDSVRKRLVDAANVNNAKDMAGKSSPLHFAAG 225  
 QY 142 FGRKDVVEYLQNCANVOARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNWYTP 201  
 Db 226 FGRKDVVEHLLQNCANVHARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNWYTP 285  
 QY 202 LHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKDELLEARS 261  
 Db 286 LHEAAIKGKIDVCIVLQHGADPNIRNTDGSALDADPSAKAVLTGEYKDELLEARS 345  
 QY 262 GNEEKMALITPLNVNCHASDGRKSTPLHLAGYNNRVKIVOLLQHGADYHAKDKGLVP 321  
 Db 346 GNEEKMALITPLNVNCHASDGRKSTPLHLAGYNNRVKIVOLLQHGADYHAKDKGLVP 405  
 QY 322 LHNACSYGHYEVELLYVKHACVNMADLWQFTPLHAEASKNRVEVCSLLLSYGADPTLLN 381  
 Db 406 LHNACSYGHYEVELLYVKHACVNMADLWQFTPLHAEASKNRVEVCSLLLSYGADPTLLN 465  
 QY 382 CHNKSALDAPTPQLKERLAYEFKSHSLQAARADYTRIKKHLSLEWFKHPQTHETA 441  
 Db 466 CHGKSADMAPTPELRERLYEFKSHSLQAARADYTRIKKHLSLEWFKHPQTHETA 525  
 QY 442 LHCAASPYKPKQICELLKRGANINEKTEFTPLHVAASEKAHNDVNVVVKHEAKVN 501  
 Db 526 LHCASVSLHPKRVQVTELLRKGANVNEKNDFMTPPLHVAARAHNDVNVVVKHEAKVN 585  
 QY 502 ALDNLGTSLHRAAYCGHLCQTRLLSYGCDPNIIISLQGTALQMGNNVQQLQEGEISL 561  
 Db 586 ALDTLGTALHRAALAGHLQTRLLSYGSDPSIISLQGTAAQMGNEAVQQLISESTPI 645  
 QY 562 GNSEADROLLEAAKAGDVTETVKKLCTVQVQNCRDIEGRQSTPLHFAAGYNNRVSVYLLQ 621  
 Db 646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNNRVSVYLLH 705  
 QY 622 HGADVHAKDKGLVPLHNACSYGHEVAELLVKHGVNVNADLWKETPLHFAAAGKYEI 681  
 Db 706 HGADVHAKDKGLVPLHNACSYGHEVAELLVKHGVNVNADLWKETPLHFAAAGKYEI 765  
 QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTIDQLLKGDAALLDAKKGCLARVKKLSSP 741  
 Db 766 CKLLQHGADPTKKNRDGNTPLDLVKDGTIDQLLKGDAALLDAKKGCLARVKKLCTP 825  
 QY 742 DNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801  
 Db 826 ENINCRDTQGRHSTPLHLAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDI 885  
 QY 802 AALLIKYNCVNAATDKWAFPLHEAAQGRQLCALLAHGADPTLKNQSGQTPDLVSA 861  
 Db 886 AALLIKYNCVNAATDKWAFPLHEAAQGRQLCALLAHGADPTLKNQSGQTPDLVSA 945  
 QY 862 DDVSAALLTAAMPSPALPSCYKQVNLNGVRSPGATADALSQSPSSLSAASSLDNLGSG 921  
 Db 946 DDIRALLIDAMPPEALPTCFKPOAT-----VVSASLISPASTPSCLSAASSIDNLGTP 998  
 QY 922 FSELSSVVSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDFIREQITLDLV 979  
 Db 999 LAELAVGASNAGDGAAGTERKECEVAGLDNLSQFLKSLGLEHLRDFITEQITLDVLA 1058





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 07:42:50 : Search time 16.9573 Seconds  
(without alignments)  
2151.551 Million cell updates/sec

Title: US-09-843-159B-4

Perfect score: 6464

Sequence: 1 RCSARRGAAGGAGQARGARV.....AYPEYLITYQIMRPGMVDG 1240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_5/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_5/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_5/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	6053	93.6	1166	4	US-09-350-982C-5
2	5103.5	79.0	1327	4	US-09-196-387-2
3	3307	51.2	949	4	US-09-196-387-10
4	2132.5	33.0	673	4	US-09-196-387-8
5	877	13.6	1839	2	US-09-172-977-4
6	858	13.3	843	2	US-09-172-977-3
7	835	12.9	1745	2	US-09-031-485-33
8	835	12.9	1745	2	US-08-847-429A-33
9	835	12.9	1745	3	US-09-065-474-33
10	835	12.9	1745	3	US-09-557-034-33
11	629.5	9.7	1088	4	US-09-082-059-2
12	514.5	8.0	352	3	US-09-065-474-139
13	514.5	8.0	352	4	US-09-557-034-139
14	421.5	6.5	303	2	US-09-031-485-23
15	421.5	6.5	303	2	US-08-847-429A-23
16	421.5	6.5	303	3	US-09-065-474-23
17	421.5	6.5	303	3	US-09-557-034-23
18	409	6.3	302	2	US-09-031-485-38
19	409	6.3	302	2	US-08-847-429A-38
20	409	6.3	302	3	US-09-065-474-38
21	409	6.3	302	4	US-09-557-034-38
22	398.5	6.2	1719	4	US-09-605-785-378
23	398.5	6.2	1719	4	US-09-439-313-378
24	398.5	6.2	1719	4	US-09-352-616A-378
25	392	6.1	348	2	US-09-031-485-28
26	392	6.1	348	2	US-08-847-429A-28
27	392	6.1	348	3	US-09-065-474-28

```

28      392      6.1      348      4      US-09-557-034-28      Sequence 28, Appl
29      390.5     6.0      787      4      US-09-188-930-334      Sequence 334, App
30      345      5.3     1423      4      US-09-810-712-10      Sequence 10, Appl
31      332.5     5.1     835      4      US-09-291-839-2       Sequence 2, Appli
32      310      4.8     191      2      US-09-031-485-20      Sequence 20, Appl
33      310      4.8     191      2      US-08-847-429A-20      Sequence 20, Appl
34      310      4.8     191      3      US-09-065-474-20      Sequence 20, Appl
35      310      4.8     191      4      US-09-557-034-20      Sequence 20, Appl
36      306      4.7     422      2      US-08-484-575A-6       Sequence 6, Appli
37      306      4.7     422      3      US-08-477-459-6       Sequence 6, Appli
38      306      4.7     422      3      US-08-479-869-6       Sequence 6, Appli
39      306      4.7     422      4      US-08-486-414-6       Sequence 6, Appli
40      306      4.7     422      5      PCT-US94-01826A-6      Sequence 6, Appli
41      306      4.7     422      5      PCT-US94-02352A-6      Sequence 6, Appli
42      305      4.7     679      2      US-08-462-481-4       Sequence 4, Appli
43      305      4.7     679      2      US-08-436-771-6       Sequence 6, Appli
44      305      4.7     679      2      US-08-434-998-6       Sequence 6, Appli
45      305      4.7     679      2      US-08-487-797-6       Sequence 6, Appli

```

#### ALIGNMENTS

```

RESULT 1
US-09-350-982C-5
: Sequence 5, Application US/09350982C
: Patent No. 6455290
: GENERAL INFORMATION:
: APPLICANT: Berthelsen, Jens
: APPLICANT: Toma, Salvatore
: APPLICANT: Isacchi, Antonella
: TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods R
: TITLE OF INVENTION: Same
: FILE REFERENCE: PHRM-0043
: CURRENT APPLICATION NUMBER: US/09/350,982C
: CURRENT FILING DATE: 1999-07-09
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 5
: LENGTH: 1166
: TYPE: PRT
: ORGANISM: Artificial
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: (1102)..(1102)
: OTHER INFORMATION: n is any nucleic acid
: NAME/KEY: misc_feature
: LOCATION: (2650)..(2650)
: OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-5

```

```

Query Match      93.6%; Score 6053; DB 4; Length 1166;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1157; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 75 MSGRRCAGGGAACASASAAAEVPAARELFECACRNGDVERVKRLVTPKVNSTRDTAGRKST 134
    |||||
Db 1 MSGRRCAGGGAACASASAAAEVPAARELFECACRNGDVERVKRLVTPKVNSTRDTAGRKST 60
    |||||
QY 135 PLHPAAGFGRKDVVEYLQNCANVQARDGGLIPLHNACSFHAEVNNLLRHGADPNAR 194
    |||
Db 61 PLHLAAGFGRKDVVEYLQNCANVQARDGGLIPLHNACSFHAEVNNLLRHGADPNAR 120
    |||
QY 195 DNWNYPPLHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLPSPSAKAVLTGEYKKDE 254
    |||||
Db 121 DNWNYPPLHEAAIKGKIDVCIVLQHGAEPTIRNTDGTALDLPSPSAKAVLTGEYKKDE 180
    |||||
QY 255 LLESARGSGNEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 314
    |||||
Db 181 LLESARGSGNEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAK 240
    |||||

```

```
QY 315 DKGDVPLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLSYG 374
Db 241 DKGDVPLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLSYG 300
QY 375 ADPTLLNCINKSAIDLAPTPQKRLAYEFKSHLSLQAAAREADVTRIKKHSLEWVNFKH 434
Db 301 ADPTLLNCINKSAIDLAPTPQKRLAYEFKSHLSLQAAAREADVTRIKKHSLEWVNFKH 360
QY 435 POTHETALHCAASAPYKPKQICELLRRKGANINEKTEFLPLHVASEKAHNDVVEVV 494
Db 361 POTHETAXHCAASAPYKPKQICELLRRKGANINEKTEFLPLHVASEKAHNDVVEVV 420
QY 495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALMGNNVQQL 554
Db 421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLSYGCDDPNIIISLOGFTALMGNNVQQL 480
QY 555 LQEGISLGNSEADRLLEAKAGADVETVKKLCTVQSVNCRDTEGROSTPLHFAAGNRYVS 614
Db 481 LQEGISLGNSEADRLLEAKAGADVETVKKLCTVQSVNCRDTEGROSTPLHFAAGNRYVS 540
QY 615 VVEYLQHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVAMDLWKFPTPLHEAA 674
Db 541 VVEYLQHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVAMDLWKFPTPLHEAA 600
QY 675 AKGYEICKLLQHGADPTKKNRDGNTPDLVKDGDPTDQDLRGDAALLDAKKGCLAR 734
Db 601 AKGYEICKLLQHGADPTKKNRDGNTPDLVKDGDPTDQDLRGDAALLDAKKGCLAR 660
QY 735 VKLSSPDVNVNCRDTRGHTSTPLHLAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 794
Db 661 VKLSSPDVNVNCRDTRGHTSTPLHLAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 720
QY 795 SYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQCALXLAHAGDPTLKNEGOT 854
Db 721 SYGHVDVAALLIKYNACVNATDKWFTPLHEAAKQRTQCALXLAHAGDPTLKNEGOT 780
QY 855 PLDLVSADVSALLTAAMPSPALPCYKQVNLGVNRPSPGATADALSSGSPSSLSAAS 914
Db 781 PLDLVSADVSALLTAAMPSPALPCYKQVNLGVNRPSPGATADALSSGSPSSLSAAS 840
QY 915 LDNLSSGFSFELSVSSSGTEGASSLEKKEVPQVDFSIQFVRLGLEHLMDFEREQIT 974
Db 841 LDNLSSGFSFELSVSSSGTEGASSLEKKEVPQVDFSIQFVRLGLEHLMDFEREQIT 900
QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPYTLTSGSGTILDLSP 1034
Db 901 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISQOGLNPYTLTSGSGTILDLSP 960
QY 1035 DKKEFOSVEEEMQSTVREHRDGHAGGIENRNYNLIKQVCNKKLWERTYTHRRKEYSEEN 1094
Db 961 DKKEFOSVEEEMQSTVREHRDGHAGGIENRNYNLIKQVCNKKLWERTYTHRRKEYSEEN 1020
QY 1095 HNHANERMLFHGSPFVNAILHKGFDERHAYIGMFCAGIYFAENSCKSNQYVYGIGGGTG 1154
Db 1021 HNHANERMLFHGSPFVNAILHKGFDERHAYIGMFCAGIYFAENSCKSNQYVYGIGGGTG 1080
QY 1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGPHSHVTRGPRSVNGLALAEYV 1214
Db 1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLOFSAMKMAHSPGPHSHVTRGPRSVNGLALAEYV 1140
QY 1215 IYRGEQAYPEYLITYOIMRPEGVMDG 1240
Db 1141 IYRGEQAYPEYLITYOIMRPEGVMDG 1166
```

## RESULT 2

US-09-196-387-2

; Sequence 2, Application US/09196387

; Patent No. 6277613

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS

```
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-196-387-2
```

## Query Match

79.0%; Score 5103.5; DB 4; Length 1327;

Best Local Similarity 79.2%; Pred. No. 0;

Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

```
QY 22 AAHTGAPDPTAGSQAARALSSPGGLALLLAGPGLLRLLALLLAVAAARIMSGRRCA 81
Db 112 SAAGVAPNAPAGSGNNSSPSSPTSSSSSSPSG-----SSLAESPEAGVSSAPL 165
QY 82 GGAACASAAAEEAVEPAARELFEACRNGDVERVKRLVTEKVNSRDTAGKSTPLHFAAG 141
Db 166 GGAAGGPGTGVPAVSGALRELLACRNGDVSVRKRLVDAANVNAKDMAGRKSPHLFAAG 225
QY 142 FGRKDVVEYLLQGANVQARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNYTP 201
Db 226 FGRKDVVEYLLQGANVQARDGGLIPLHNACSFHAEVYNLLLRHGADPNARDNNYTP 205
QY 202 LHEAAIKGKIDVCIVLLQHGAEPTINTDGTALDLADPSAKAVLTGEYKKDELLESARS 261
Db 286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGSALDLADPSAKAVLTGEYKKDELLEAARS 345
QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRVKIVQLLQHGADVHAKDKGLVP 321
Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAGYNRVKIVQLLQHGADVHAKDKGLVP 405
QY 322 LHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLSYGADPTLN 381
Db 406 LHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSLLSYGADPTLN 465
QY 382 CHNKSIDLAAPTQKRLAYEFKSHLSLQAAAREADVTRIKKHSLEWVNFKPQTHETA 441
Db 466 CHNKSIDLAAPTQKRLAYEFKSHLSLQAAAREADVTRIKKHSLEWVNFKPQTHETA 525
QY 442 LHCAASPYKPKQICELLRRKGANINEKTEFLPLHVASEKAHNDVVEVYVYKHAQVN 501
```

Db 526 LHCASVLPKPKQVTELLRRKANVNEKNKDFMTPLHVAABRAHNDVNEVLKKGAKMN 585  
Qy 502 ALDNLGQTSLHRAAYCGHLOTCRLLSYGCCPNITISLOGFTALOMGNENVOQLLOEGTSL 561  
Db 586 ALDTLGQTLHRAALAGHLQTCRLLSYGSDPSIISLOGFTAAQMGNEAVQOILSESTPI 645  
Qy 562 GNSEADROLLEAAKAGDVETVKKLTQVSVCNRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621  
Db 646 RTSVDYRLLEASKAGDLETKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705  
Qy 622 HGADVHAKDKGLVPLHNAACSYGHVEAEVLLVKGAVNVADLWFTPLHFAAAKGYEI 681  
Db 706 HGADVHAKDKGLVPLHNAACSYGHVEAEVLLVKGAVNVADLWFTPLHFAAAKGYEI 765  
Qy 682 CKLLHOGADPTKKNROGNTPLDLVKDGTDTQDLRLGDAALLDAKKGCLARVKKLSPP 741  
Db 766 CKLLHKGADPTKKNROGNTPLDLVKDGTDTQDLRLGDAALLDAKKGCLARVKKLSPP 825  
Qy 742 DNVNCRDTQGRHSTPLHNAAGYNNLEVAEYLLHOGADYNAQDKGGLIPLHNAASYGHVDV 801  
Db 826 ENINCRDTQGRNSTPLHNAAGYNNLEVAEYLLHOGADYNAQDKGGLIPLHNAASYGHVDI 885  
Qy 802 AALLIKYNACYNATDKWFTPLHNAAGYNNLEVAEYLLHOGADYNAQDKGGLIPLHNAASYGHVDI 861  
Db 886 AALLIKYNACYNATDKWFTPLHNAAGYNNLEVAEYLLHOGADYNAQDKGGLIPLHNAASYGHVDI 945  
Qy 862 DVSALLTAAMPSPALPCYKQVNLGVKSPGATADALSSGSPSSLSAASSLDNLSSS 921  
Db 946 DDIRALLIDAMPPEALPCFKPQAT-----VVSASLISPASTPSCLSAASSLDNLGTP 998  
Qy 922 FSELSSVVSSGTGASLEKK--EVPGVDFSITQFVRNLGLEHLMIDIFEREQITLDVLV 979  
Db 998 LAELAVGASNAGDGAAGTERKEGEVAGLDNMIISOFLKSLGLEHLRDIETEQITLDVLA 1058  
Qy 980 BMGHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPVLTNTSGSTILDLSPDDKEF 1039  
Db 1058 DMGHEELKEIGINAYGHRHKLKIGVERLISGOQGLNPVLTNTSGSTILDLSPDDKEF 1118  
Qy 1040 QSVBEEMQSTVREHRDGGHAGGIFNRNLIKIQVCNKKLWERYTHRRKEYSEENHNAN 1099  
Db 1118 QSVBEEMQSTVREHRDGGHAGGIFNRNLIKIQVCNKKLWERYTHRRKEYSEENHNAN 1178  
Qy 1100 ERMFLHSGSPFNALIHKGFDERHAYIGMFCAGIYFAENSCKSNQYVYIGGGTGCPCVHK 1159  
Db 1178 ERMFLHSGSPFNALIHKGFDERHAYIGMFCAGIYFAENSCKSNQYVYIGGGTGCPCVHK 1238  
Qy 1160 DRSCYICHRQLFCRVTLGKSLQFSAMKMAHSPGHHSVTCGRSVNGLALAEYVIRGE 1219  
Db 1238 DRSCYICHRQLFCRVTLGKSLQFSAMKMAHSPGHHSVTCGRSVNGLALAEYVIRGE 1298  
Qy 1220 QAYPEYLITYQIMRPE 1235  
Db 1299 QAYPEYLITYQIMRPE 1314

RESULT 3  
US-09-196-387-10  
; Sequence 10, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; City: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196,387  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/095,225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 949 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-196-387-10  
Query Match 51.2%; Score 3307; DB 4; Length 949;  
Best Local Similarity 76.2%; Pred. No. 6,4e-279;  
Matches 638; Conservative 72; Mismatches 119; Indels 8; Gaps 2;  
Qy 22 AAHGTPAPVPTAGSOAARLASASSPGGLIALLAGPGLLRLLRLLALLAVAAARIMSGRCA 81  
Db 112 SAAGVAPAPAGSGNNSSSSSSPTSSSSSSSPG-----SSLAESPEAAGVSSTAPL 165  
Qy 82 GGAACAASAAAEVAPFAARFELFEACRNGDVERVKRLVTPEKYNRDTAGRKSTPLHFAAG 141  
Db 166 GPGAAGPGTGPVAVSALRELLEACRNGDVSRLVDAANVNAKMAGRKSSPLHFAAG 225  
Qy 142 FGRKDVVEYLLONGANVARDGGGLIPLHNAACSFHAEVYVNLRLRHGADPNARDNNWYTP 201  
Db 226 FGRKDVVEYLLONGANVARDGGGLIPLHNAACSFHAEVYVNLRLRHGADPNARDNNWYTP 285  
Qy 202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDLPASAKAVLTGEYKKDELLESARS 261  
Db 286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGSALDLPASAKAVLTGEYKKDELLESARS 345  
Qy 262 GNEEKMAALLPLNVNCHASDGRKSTPLHAAAGYNRVIVQLLQHGADVHAKDKGLVLP 321  
Db 346 GNEEKMAALLPLNVNCHASDGRKSTPLHAAAGYNRVIVQLLQHGADVHAKDKGLVLP 405  
Qy 322 LHNACSYGHVEVTELLVKGACVNAWDLWQFTPLHAAASKNRVYCSLLLSYGADPTLLN 381  
Db 406 LHNACSYGHVEVTELLVKGACVNAWDLWQFTPLHAAASKNRVYCSLLLSYGADPTLLN 465  
Qy 382 CHNKAIDLAPTOLKERLAYEFKGHSLQAAREADVTIRKKHLSLEVMNEKHPOTHETA 441  
Db 466 CHGKSAVDMATPELRERLTVEFKHSLQAAREADLAKVKKTLALEINFKQPOSHETA 525  
Qy 442 LHCAASVYPKPKQVTELLRRKANVNEKNKDFMTPLHVAABRAHNDVNEVLKKGAKMN 501  
Db 526 LHCAASVYPKPKQVTELLRRKANVNEKNKDFMTPLHVAABRAHNDVNEVLKKGAKMN 585  
Qy 502 ALDNLGQTSLHRAAYCGHLOTCRLLSYGCCPNITISLOGFTALOMGNENVOQLLOEGTSL 561  
Db 586 ALDTLGQTLHRAALAGHLQTCRLLSYGSDPSIISLOGFTAAQMGNEAVQOILSESTPI 645  
Qy 562 GNSEADROLLEAAKAGDVETVKKLTQVSVCNRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621  
Db 646 RTSVDYRLLEASKAGDLETKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705  
Qy 622 HGADVHAKDKGLVPLHNAACSYGHVEAEVLLVKGAVNVADLWFTPLHFAAAKGYEI 681

Db 706 HGADVHAKDKGGLVPLHNACSYGHVEAEVLLVRHGASVNVADLWKFTPLHEAAAKGYEI 765  
QY 682 CKLLQHGADPTKKNRDNTPDLVYKDGTDIDODLLRGDAALLDAAKKCLARVKKLSSP 741  
Db 766 CKLLKKGADPTKKNRDNTPDLVYKDGTDIDODLLKGDAAALLDAAKKGLARVQKLCPT 825  
QY 742 DVNCRDGTGRSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPLPHNAASYGH--V 799  
Db 826 ENINCRDGTGRSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPLPHNAASYGGCLA 885  
QY 800 DVAALLIKYNACVNAIDKWAFTPLHEAAQGRTOQLCALLLAHGADPTLKNQEGTPL 856  
Db 886 RVQKCTPENINCRODQGRNSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLPL 942

RESULT 4  
US-09-196-387-8  
; Sequence 8, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196,387  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/095,225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIPI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 673 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-196-387-8

Query Match 33.0%; Score 2132.5; DB 4; Length 673;  
Best Local Similarity 73.3%; Pred. No. 7.6e-177;  
Matches 417; Conservative 54; Mismatches 91; Indels 7; Gaps 2;

QY 22 AANGTAPDPTAGSQARALSSPGGLALLLAGPGLILRLALLLAVAAARIMSGRCA 81  
Db 112 SAAGVAPNPAGSGNSPSSSSPTSSSSSPSPG-----SSLAESPEAAGVSTAPL 165  
QY 82 GGCAACASAAAEVAPAELEFACRNGDVERVKRLVTPPEKVNSRDTAGRKSTPLHFAAG 141  
Db 166 GPGAGGPGTGPVAVSGALRELEACRNGDVSRYKRLVDAANVNAKMGAKSSPLHFAAG 225  
QY 142 FGRKDVVEYLLQNCANVQARDGGLPLHNACSFHGAEEVNVNLLLRHGADPNARDNNYTP 201

Db 226 FGRKDVVEYLLQNCANVHARDGGLPLHNACSFHGAEEVNVNLLLRHGADPNARDNNYTP 285  
QY 202 LHEAAIKGKIDVICIVLLQHGAEPTIRNTDGRDALDADPSAKAVLTGEYKKDBELLESARS 261  
Db 286 LHEAAIKGKIDVICIVLLQHGADPNIRNTDGRSALDADPSAKAVLTGEYKKDBELLESARS 345  
QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVKIVQVLLQHGADVHAKDKGLVP 321  
Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVKIVQVLLQHGADVHAKDKGLVP 405  
QY 322 LHNACSYGHEVTELLVYKHGACVNAWDLWQFTPLHFAASKNRVEVCSLLSYGADPTLN 381  
Db 406 LHNACSYGHEVTELLVYKHGACVNAWDLWQFTPLHFAASKNRVEVCSLLSYGADPTLN 465  
QY 382 CHNKSALDAPLPOLKERLAYEFKSHLSLQAAAREADVTRIKKHLSLNMFVNFKHPQTHETA 441  
Db 466 CHKSASVDMAPTPELRERLTYEFKSHLSLQAAAREADVTRIKKHLSLNMFVNFKHPQTHETA 525  
QY 442 LHCAASPYPKRKQICELLRLKGANINEKTKFPLTHVASEKAHNDVVEVYVYKHEAKVN 501  
Db 526 LHCAVASLHPKRQVTELLRLKGANVNEKNKDFMTPLHVAARAHNDVMEVLKHGAKMN 585  
QY 502 ALDNLGQTSLHRAAYCGHLQTCRLLSYCGDPMIISLQGFALOMGNVQVLLQEGISL 561  
Db 586 ALDNLGQTSALHRAALAGHLQTCRLLSYSGSDPSIISLQGFALOMGNVQVLLSYVSGS 645  
QY 562 GNSEADROLLEAAKAGDVETVKKLCTVQS 590  
Db 646 DFIISLQGFALOMGN-EAVQVLLSGHS 673

RESULT 5  
US-09-172-977-4  
; Sequence 4, Application US/09172977  
; Patent No. 5989863  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
; FILE REFERENCE: PF-0615 US  
; CURRENT APPLICATION NUMBER: US/09/172,977  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 4  
; LENGTH: 1839  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: g29491  
US-09-172-977-4

Query Match 13.6%; Score 877; DB 2; Length 1839;  
Best Local Similarity 29.7%; Pred. No. 1.1e-66;  
Matches 293; Conservative 137; Mismatches 402; Indels 156; Gaps 23;

QY 105 ACRNGDVER-VKRLVTPPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNCANVQARD 163  
Db 38 AARAGNLDKVEYLLKGGIDINTCQNGLNA--LHAAKEGHVGLVQELLGRGSSVDSATK 95  
QY 164 GGLTPLHNACSFHGAEEVNVNLLLRHGADPNARDNNYTPPLHEAAIKGKIDVICIVLLQHGAE 223  
Db 96 KGNTALHIALSLAQAEVVKVVLKREGANINAQSQNGFTPLYMAQENHIDVVKVLLNGAN 155  
QY 224 PTIRNTDGR-----ALDADPSAKAVLTGEYKKDEL-----LESARSGNEKMMALLPLN 275  
Db 156 QSTATEDGFTPLAVALQQGHQNAVAIILENDTKGKVRPLALHIAARKDDTKSAALLQ-- 213  
QY 276 VNCHASDGRKS-----TPLHLAAGYNNRVKIVQVLLQHGADVHAKDKGLVPLHN 324



```
Db 214 -NDHNADVQSKMMVNRNTTESGFTPLHIAAHYGNVNVATLLNRCGAADVFTARNGITPLHV 272
QY 325 ACSYGHYEVTLLVKGHCACVNDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db 273 ASKRGNTNMVKLLDRGGQIDAKTRDGLTPLHCAARSHGQDVVELLGERGA-PLIARTKN 331
QY 385 KSAIDLAPTOLKRLAYEFKHSLLQAAREADVTRIKKHLISLEMVNFKHP-----QTHE 439
Db 332 -----GLSPLHMAAGDHVECVKHL-----LOHKAPVDVDTLDYL 366
QY 440 TALHCAASPYPRKQICELLLRKGANINEKTEFLPLHVASBKANDVVEVVKHAK 499
Db 367 TALHVA-----HCGHYRVTKLLDKRANPNARALNGFTPLHIACKNRIKVMELLYKYGAS 423
QY 500 VNALDNLGOTSLHRAAYCGHLCOTCRLLLSYCCDPNIISLOGETALOMGN-----ENVQOL 554
Db 424 IQAITESGLTPIHVAAPFNGHLNIVLLLLONGASPDVNIIRGETALHMAARAGQVEVVRCL 483
QY 555 LOEGI-----SLGNSEADROLLE-----AAKAGDV 579
Db 484 LRNGALVDARAREBOTPLHIASRLGKTEIVQLLQHHMAHPDAATTNGYTPHISAREQV 543
QY 580 ETYKKTCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
Db 544 DVASVLEAGAAHSLATK-RGFTPLHVAAYKGSLDVAKLLQRRRAADSAGKNGLTPLHV 602
QY 640 ACSYGHYEVAELLVKHGAVNVADLWKFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 699
Db 603 AAHYDNQKVALLEKASPHATAKNCTYTPHIAAKKNQMCIASTLLNYGAETNIVTKQG 662
QY 700 NTPDLV-KGDDTDIOLDRGDAALDAKGLCLARYKLSLSPDNVNCRDTOGRHSTPLH 758
Db 663 VTPLHLASQEGHDMVTLLDKGANIMSTKSGL-----TSLH 700
QY 759 LAAGYNNEVAEYLLQHGADVNAQDKGLIPLHNAASGYHVDVAALLIKYNACVNATDKW 818
Db 701 LAAGKVNVDAILTHGADQDAHTKLGTYPLIVACHYGNVMKVNFLLLKOGANVNAKTN 760
QY 819 AFTPLHEAAGKRTOLCALLAHGADPTLKNQEGQTPDL-----VSADDVSALLTAAM 872
Db 761 GYTPPLHQAQOQTHIINVLLQHGAKPNATTANGNTALATAKRLGYSVVDTLKVYTEEV 820
QY 873 PPSALPSCYKQVINGVRSPGATADALSSGSPSSLSAASSLDNLSGSPSEL-----925
Db 821 TTTTITITEK-----HKLNPETMTEVL-----DVSDEEGDDTMTGCGEYLRPEDLK 868
QY 926 -----SSVSVSSSGTEGASSLEKKEVPGVDFESITQFV--RNLGLEHLMDFEREQITLDVLV 979
Db 869 ELGDDSLPSSQFLDGMNYLRYSLGGRSDSLRFSFSDRSHTLSHAS--YLROSAVMDDSV 926
QY 980 EMGHKELKEIGINAYGHRHKLKIGVERL 1007
Db 927 VIPSHQVSTLAKEARNYSYRLSWGTEL 954
```

## RESULT 6

```
US-09-172-977-3
; Sequence 3, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PR1
```

```
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: g1841966
US-09-172-977-3
```

Query Match 13.3%; Score 858; DB 2; Length 843;

Best Local Similarity 29.6%; Pred. No. 1.3e-65;

Matches 275; Conservative 148; Mismatches 367; Indels 140; Gaps 23;

```
QY 105 ACRNGDVER-VKRLVTPKVNRSRTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARD 163
Db 3 AARAGNLDKVVYELKGGIDINTCQNGUNA--LHLAAKEGHVGLVQELLGRSSVDSATK 60
QY 164 GGLIPLHNACSFHAEVYVNLRLRHGADPNARDNNYTPPLHEAAATKGDVCIIVLLOHAE 223
Db 61 KGNTALHIASLAGAEVYKVLVKEGANINAQSONGFTPLYMAAOGHHIDVYKYLLENGAN 120
QY 224 PTIRTDGRT-----ALDLADPSAKAVLTGEYKKDEL-----LESARSGNEEKMMALLTPLN 275
Db 121 QSTATEDGFTPLVALOQGHNOAVAILLENDTKGVRLPALHIAARKDDTTSAAALLQ-- 178
QY 276 VNCHASGRKS-----TPHLAAGYNRVKIVOLLQHGADVHAKDKGDLVPLHN 324
Db 179 -NDHNADVQSKMMVNRNTTESGFTPLHIAAHYGNVNVATLLNRCGAADVFTARNGITPLHV 237
QY 325 ACSYGHYEVTLLVKGHCACVNDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db 238 ASKRGNTNMVKLLDRGGQIDAKTRDGLTPLHCAARSHGQDVVELLGERGA-PLIARTKN 296
QY 385 KSAIDLAPTOLKRLAYEFKHSLLQAAREADVTRIKKHLISLEMVNFKHP-----QTHE 439
Db 297 -----GLSPLHMAAGDHVECVKHL-----LOHKAPVDVDTLDYL 331
QY 440 TALHCAASPYPRKQICELLLRKGANINEKTEFLPLHVASBKANDVVEVVKHAK 499
Db 332 TALHVA-----HCGHYRVTKLLDKRANPNARALNGFTPLHIACKNRIKVMELLYKYGAY 388
QY 500 VNALDNLGOTSLHRAAYCGHLCOTCRLLLSYCCDPNIISLOGETALOMGNVQQLQEGI 559
Db 389 IQAITESGLTPIHVAAPFNGHLNIVLLLLONGASPDVNIIRGETALHM-----435
QY 560 SLGNSEADROLLEAAKAGDVETVKKCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
Db 436 -----AARAGEVEVVR--CLLRNGALVDARAREEQTPHLIASRLGTEIVOL 480
QY 619 LLOHGADVHAKDKGGLVPLHNAASGYHVEAELLVKHGAVNVADLWKFTPLHEAAGK 678
Db 481 LLOHMAHPDAATTNGYTPHISAREGOVDVASVLEAAGAAHSLATKKGFTPLHVAAYKS 540
QY 679 YEICKLLOHGADPTKKNRDGNTPLDLVKDGDITDIQDLRGDAALLDAAKGCLARYKKL 738
Db 541 LDVAKLLQRRRAADSAGKNGLTPLHVAAYDN-----QKVALLLLEK 584
QY 739 SSPDNVNCRDTOGRHSTPLHIAAGYNNEVAEYLLQHGADVNAQDKGLIPLHNAASYGH 798
Db 585 ASP-----HATAKNGYTPHIAAKKNQMCIASTLLNYGAETNIVTKGVTPHLASQEGH 639
QY 799 VDVAAALLIKYNACVNATDKWAFTPHIAAGKRTOLCALLAHGADPTLKNQEGQTPDL 858
Db 640 TDMVTLVLEKGANIHMSTKSGLTSLHAAEEDKYNVADILTKHGADQADATKLGTYTPLIV 699
QY 859 V-----SADDVSALLTAAMPESA-LPSCYKP-----QVLNGVRSFGATDALSSGP 903
Db 700 ACHYGNVMKVNFLLLKOGANVNAKTNGYTPHQAQOQTHIINVLLQHGAKPNATTA-- 757
QY 904 SSPSSLSAASSLDNLSGFSSELSVSSSGTEGASSLEKKEVPGVDFESITQFVRNLGLEH 963
Db 758 NGNTALAIARLGYI--SVVDTLKVVTTEVTTTTTITEKHKLNAPEMTIE-----806
QY 964 LMDIFERE---QITLDVLVEMGHKELKEIG 990
Db 807 VLDVSDDEEGDVTVTGDGGEYLRPDLKELG 836
```

RESULT 7  
 US-09-031-485-33  
 ; Sequence 33, Application US/09031485  
 ; Patent No. 5824306  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Liang  
 ; APPLICANT: Blehm, E. Scot  
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
 ; TITLE OF INVENTION: USES THEREOF  
 ; NUMBER OF SEQUENCES: 85  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.  
 ; ADDRESS: Heskia Corporation  
 ; STREET: 1825 Sharp Point Drive  
 ; CITY: Fort Collins  
 ; STATE: Colorado  
 ; COUNTRY: USA  
 ; ZIP: 80525  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: Windows 95  
 ; SOFTWARE: WordPerfect for Windows, Version 7.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/031.485  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/847.429  
 ; FILING DATE: 24-APR-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Verser, Carol Talkington  
 ; REGISTRATION NUMBER: 37,459  
 ; REFERENCE/DOCKET NUMBER: HW-5  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 970/493-7272  
 ; TELEFAX: 970/484-9505  
 ; INFORMATION FOR SEQ ID NO: 33:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1745 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-031-485-33

Query Match 12.9%; Score 835; DB 2; Length 1745;  
 Best Local Similarity 25.7%; Pred. No. 4.5e-63;  
 Matches 314; Conservative 154; Mismatches 452; Indels 304; Gaps 37;  
 QY 34 GSOAARALSASSPGGLALLAGPGLLLRLALLAVAAARIMSG---RR-AGGGAACAS 89  
 DB 35 GESSASFRAARAGNLD-----RVLELL-----RSGTDINTCNANGNALHL 76  
 QY 90 AAAAEVAPAAARELFACRNGDVERVKRLVTPBKVNSRDITAGRK-STPLHFAAGFGKDVV 148  
 DB 77 ASKEGHVHVRELLK--RKADV-----DAATRKGNLTALHSLAQOELIV 119  
 QY 149 EYLLONGANVQARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDWNVYTPLEAAIK 208  
 DB 120 TVLVENGANVNVQSLNGFTFLYMAAQENHESVRYLLAHNAQAALSTEDGFTPLAVALQ 179  
 QY 209 GKIDCVILVHGAETIRNTDGRALDIA-----DPSAKAVLTGETYKKDELLES----- 258  
 DB 180 GHDVVAVULENDTRGKVR-----LPALHTAAKDDTKAATLLQNEHNSDVTSKSGFTPL 235  
 QY 259 ---ARSGNEEKMMALL-----TPLNV-----NCHAS 281  
 DB 236 HIAAHYGNENVAQLLLEKGANVYQARHNISPLHVATKWRGTMVSVLLLAHGAVIDCRTR 295  
 QY 282 DGRKSTPLHAGYKNVYQVLLQHGADVHAKDKGLVPLHNACSY----- 328

DB 296 D--LITPLHCASRSGHDQVVDLLLEKGAPISAKTKNGKLAPLHMAAQVDDVTVDTPLHV 353  
 QY 329 ---CHYEVTLLYKHGACVNMADLWQFTPLHAEASKNRVVCSSLISYGADPTLLNCHN 384  
 DB 354 AAHCHVRAKLLLDNRADPNARALNGFTPLHIAKKNRIKIVELLKYHA----- 404  
 QY 385 KSAIDLAPTOLKERLAYEFKGH-----SLQAAAREADVTRIKHLSLEMYNFKHPOTHE 439  
 DB 405 --ATEATTESGLSPLHVAAPFAGINIVYLLQOGANADVATVR-----GE 447  
 QY 440 TALHCAASPYPKRQICELLIRKGANINEKTELTPLHVASEKAHNDVVEVVYKHEAK 499  
 DB 448 TPLHLAARA---NOTDIVRVLRNGAQVDAARELQTPHLHIAASRLGNTDIVILLQANAS 504  
 QY 500 VNALDNLGQTSILHRAAYCGHLQTCRLLSYCGDPMIISLQGFALQMGNNENVQQLQEG 559  
 DB 505 PNAATRDLYTPLHIAAKEGQEVAAIIMDHGTDKTLTKKGFPLHL----- 551  
 QY 560 SLGNSEADROLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618  
 DB 552 -----AAKYGNLPVAKSL--LERGTVPVDIEGKNQVTPHVAAHYNNDKVALL 596  
 QY 619 LLOHCADVHAKDKGLVPLHNACSYGHEVAELLYKHGAVNVNADLWKFPTPLHAEAAK 678  
 DB 597 LLENGASAHAAKNGYTPHIAAKKNQMDIASTLLHYKANAANAEKAGFTPLHLAAQEGH 656  
 QY 679 YEICKLLLOHGADPTKKNRDNTPDLVLDKDGDTDQDILLRGDAALLDAKKGCLARVKL 738  
 DB 657 REMAALLIENGAKVGAQARNG----- 677  
 QY 739 SSPDNVNCROTQGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 798  
 DB 678 -----LTPMHLCAQEDRVSAEELVKENAAIDPKTKAGYTPHLHVACHFQ 722  
 QY 799 VDVAALLIKYNACVNATDKWFTPLHAEAAKGRTOCALLLAHGADPTLKNQEGOTPLDL 858  
 DB 723 INNVRELIHGHARVSVITRASITPLHQAQQGHNSVRYLLEHGASPNVHTSTGTPLSI 782  
 QY 859 VS---ADDVSALLT-----AAMPSPALPCYKPOVLNGVRSFGATADAL---SSGPS 905  
 DB 783 AERLGVSVVEALKTITETVTITETVTETERYKQ-----NPEAMNETMFSDEGE 836  
 QY 906 PSSLSAASLDNLSGFSSELSSVSSG-----TEGASSLEKKEVPYGVDFSTITQVRNL 959  
 DB 837 DNOITANAHAHDFSESLTK--GLHDSGTGVLHIAHTEPLTS-RSPEVEGTGDDLDALIRKA 893  
 QY 960 GUEHL-----MDIFEREQITLDVLMGHKELKEIGINAYGH-----RHKLKNGVER 1006  
 DB 894 QHEPITTAMADPSLDASLPDNTVI-MRTMQPSFLISFVMDARGGAMGCRH---SGVRI 949  
 QY 1007 LISGQOGLNP-YLTINTSGGTI-----LIDLSRDDKQFQSV----- 1042  
 DB 950 IIPPKAKQPTRVTCRYLGKDKLAHPPLSEGEALASRIEMAPHGAKFLGPVILEVPHF 1009  
 QY 1043 ----BEEMQSTVREHDDGHAGIFNRYNLLKQCNKMLWERYTHRRKEVSEENHHA 1098  
 DB 1010 ASLRGREREIVLRSDDGQH----WKEHQLEATEDAVQEVLMNESFDEAESQLDLDLHSTR 1065  
 QY 1099 NERMLFHGSPFVNAITHKGFDERH 1122  
 DB 1066 ITRILTNDPFMTFVAVTVRVRQEVH 1089  
 RESULT 8  
 US-08-847-429A-33  
 ; Sequence 33, Application US/08847429A  
 ; Patent No. 5827692  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Liang  
 ; APPLICANT: Blehm, E. Scot  
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
 ; TITLE OF INVENTION: USES THEREOF

```

:
: NUMBER OF SEQUENCES: 85
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carol Talkington Verser, Ph.D.
: ADDRESSEE: Heska Corporation
: STREET: 1825 Sharp Point Drive
: CITY: Fort Collins
: STATE: Colorado
: COUNTRY: USA
: ZIP: 80525
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: WordPerfect for Windows, Version 7.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/847,429A
: FILING DATE: 24-APR-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Verser, Carol Talkington
: REGISTRATION NUMBER: 37,459
: REFERENCE/DOCKET NUMBER: HW-5
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 970/493-7272
: TELEFAX: 970/484-9505
: INFORMATION FOR SEQ ID NO: 33:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1745 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-847,429A-33
:
: Query Match 12.9%; Score 835; DB 2; Length 1745;
: Best Local Similarity 25.7%; Pred. No. 4.5e-63;
: Matches 314; Conservative 154; Mismatches 452; Indels 304; Gaps
:
QY 34 GSOAARALSASPPGGLALLAGPGLLRLALLLAVAAARTMSG---RRC-AGGGAACAS 89
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 35 GESSASFLLRAARGNLD-----RVLELL-----RSGTDTINTCANGLNALHL 76
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 90 AAAAEVPEAARELFTACRNGDVERVKRLVTPKVNSRDTAGRK-STPLHFAAGFGKDDV 148
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 77 ASKEGHVVRLLK--RKADV-----DAAIRKGNATLHIASLAGQELIV 119
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 149 EYLLQNGANVQARDGGGLIPLHNACSFGHAEVYNLLLRHGADPNARDNNYTPLEHAAIK 208
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 120 TVLVGNGANNVQSLNGFTPLMAAQENHESVVRYLLAHNAAQALSTEDGGTPLAVALQQ 179
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 209 GKIDVICVILHGCABETIRNTDGRALDLA-----DPSKAVLTGEYKKDELLES- 258
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 180 GHDVRVAVILENDTRGKVR----LPALHIAAKKDDTKAATLLQLQNEHNSDVTSGSGFTPL 235
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 259 ---ARSGNEKEMALL-----TPLNV-----NCHAS 281
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 236 HTAAHYGNENVAQLLEKGANVYQARHNISPLHVATKGRTNMVSLLLAHCAVIDCTR 295
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 282 DGRKSTPLHLAAGYNRKVTIQLLHGGADVHAADKGDVLPVHNACSY----- 328
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 296 D--LLTPLHCASRSGHDQVVDVLLLEKGAPISAKTKNGLAPLHMAAQVDDVTVDYLTPLHV 353
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 329 ---GHVEYTELLVKHCACVNAMDLPWFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 354 AAHCGHVRVAKLLLDNRNADPNRALNGFTPLHIACKNRKIKIVELLKYHA----- 404
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 385 KSAIDLAPTQPKERLAYEFKGH-----SLQAAREADVTRIKKLSLEMVNFKHPQTHE 439
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 405 --AIEATTESGLSPLHVAAPMGAINIVILLQGANADVATVR-----GE 447
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 440 TALHCAAAASPYPKRKQICELLIRKKGANINKEKTFETPLPHVASEKAHNDVVEVVKHEAK 499
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 448 TPLHLAARA---NOTDIDVRVLVNGAQQVDAARARELOTPLHIASRLGNTDITVILLQANAS 504
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

QY	500	VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDDPNIISLOGFTALQMGNENWQQLOEGI	559
Db	505	PNAATRDLYTPLHTAAKEGOEEVAAIIMDRHGDTKTLTKKGFTPLHL-----	551
QY	560	SIGNSEADROLLEAARAGDVETVKKLCTGVSNCRDIETR--QSTPLHFAAGYNRVSVVEY	618
Db	552	-----AAKYGNLPVAKS!----LBERGTPVDIEGKNQVTPLHVAHYNNDKVALL	596
QY	619	LLQHGDVHAOKGGVLPLHNACSYGHYEVAELLVKHGAVVNVADLMWKFTPLHEAAAKGK	678
Db	597	LLENGASAAAAANGYTPLHIAAKKNOMDIASTLLHYKANANAESKAFTPLHLAAOEGH	656
QY	679	YEICKLLLHQHADPTKKNRGNTPDLVLKDGDTDIDILLRGDAALLDAAKKGCLARYKKL	738
Db	657	REMAALLTEMGAKVGAQRNG-----	677
QY	739	SSPDNVNCRTOGRHSHTPLHLAAGYNNLEVAEYLLQHGADVNAODKGGLTPLHNAASYGH	798
Db	678	-----LTPMHLCAQEDRSVAAEELVKENAAIDPKTKAGYTPLHVACHFGQ	722
QY	799	VDVAALLIKYNACVNATDKWAFTPLHEAAOKGRFOLCALLLAHCADPTLKNQEGOTPLDL	858
Db	723	INMVRFUEHGARSVITRASYPPLHQAAGQHNSVVPYLEHGASPNVHTSGTGPLSI	782
QY	859	VS---ADDVSALLT-----AAMPSPALPSICYKPQLVNGVRSPGATADAL---SSGPSS	905
Db	783	AERLGYSVVVEALKTITETTIVITTEERYKPQ-----NPEAMNETWFSDESEGE	836
QY	906	PSSLASAASLDLNJSGSFESELSSVVSSG-----TEGASSLEKKEVCPVDIFSITQFVRL	959
Db	837	DNOITANAHADFSESUTK---GLHDSTGVLIHATEPTLS-RSPSEVGTDGDLDALIRKA	893
QY	960	GLEHL-----MDIFEROBITLDLVEMHGKKELKEIGINAYGH-----RHKLKIGVER	1006
Db	894	QHEPIITAMADPSLDASLPDNVTI-MRTTHQPSPLSIFMWVDARGAMRGCH---SGVRI	949
QY	1007	LISGQOGLNP-YUFLNTSGSTI-----LIDLSPDKKEFSQV-----	1042
Db	950	IIPPRKAPOPTRVCRYLHGDKLHAHPPLPSEGEALASRILEMAPHGAKFLGPVILEPHF	1009
QY	1043	-----EEMOSTVREHRDGGHAGGIFFNYNILKIQVCNKKLWEYTHRRKEVSENHNHA	1098
Db	1010	ASLRGREIREILSRSDGQH----WKEHOLEATEDAVOEVLNESFDAELSOLDLHTSR	1065
QY	1099	NERMLFHGSPFVNAIIHKGFDERH	1122
Db	1066	ITRLTNDFPMYFAVTVTRVOEVH	1089

```

1  RESULT 9
2  US-09-065-474-33
3  ; Sequence 33, Application US/09065474
4  ; Patent No. 6063599
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Tang, Liang
7  ; APPLICANT: Blehm, E. Scot
8  ; TITLE OF INVENTION: DROFILARIA AND BRUGIA ANKYRIN
9  ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
10 ; TITLE OF INVENTION: USES THEREOF
11 ; NUMBER OF SEQUENCES: 171
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
14 ; ADDRESSEE: Heska Corporation
15 ; STREET: 1825 Sharp Point Drive
16 ; CITY: Fort Collins
17 ; STATE: Colorado
18 ; COUNTRY: USA
19 ; ZIP: 80525
20 ; COMPUTER READABLE FORM:
21 ; MEDIUM TYPE: Floppy disk
22 ; COMPUTER: IBM PC compatible
23 ; OPERATING SYSTEM: Windows 95
24 ; SOFTWARE: WordPerfect for Windows, Version 7.0

```

TELEFAX: 970/484-9505

```

657 REMAALLIENCAKVGQAQRNG-----677
739 SSPDNVNCRTQGRHSPTPLHLAAGYNNLEVAEYLLQHDGADVNAQDKGLIPLHNAASYGH 758
678 -----LTPMHLCAQEDRVSVABELVVENAAIDPKTKAGTYTPLHVACHFCQ 722
799 VDVAAALIKYNACVNAITDKWAFTPLHEAAQKGRQTQCALLAHGADPTLKNQBGOTPLDL 858
723 INMVRFLIEHGARVSVITRASVYTPLHQAAQOGHNSVVVYLLLEHGASPNVHTSTGOTPLSI 782
859 VS-----ADDSVALLT-----AAMPSSALPSCYKPPQVLNVRSPGATADAL-----SSGPSS 905
783 AERLGYVSVVEALKTITETVITETTVTEERYKQ-----NPEAMNETMFSDESEGE 836
906 PSSLSAASSLDNLGSGFSELSSVVSSG-----TECASSLEKKEYGVDFTSITQFVRNL 959
837 DNQITANAHAHDFSESUTK--GLHSDTGVHLIHATETPLS-RSEVEGTGDGDLALIRKA 893
960 GLEHL-----MDIPERQITDLVLVEMGHKELKEIGINAYGH-----RHKLIKGVGR 100
894 QHEPITTMADPSLDASLPDNVIL-MRTTQPSFELISFMVDARGGAMRCRH---SGVRI 949
1007 LISQOQGLNP-YITLNVSSGSI-----LIDLSPDKKEFQSV-----104
950 IIPPRKAPQTRTVCRVLGRDKLAHPPLSEGEALASRIEMAPHGAKFLGPVILEVPHF 100
1043 -----EEEMOSTVREHRDGGHAGGFNRYNITLKTQVCNKKLWERYTHRRKEYSEENHNIA 109
1010 ASLURGREGREIVILRSSDQGH-----WKEHQLEATEDAVQEVLSNFSDAEELSQDLDLHTSR 106
1099 NERMLFHGSPFVNAILHKGFDERH 1122
1066 ITRLLINDFFMYFAVTVRQEVH 1089

RESULT 10
US-09-557-034-33
: Sequence 33, Application US/09557034
: Patent No. 6365569
: GENERAL INFORMATION:
: APPLICANT: Tang, Liang
: Blehm, E. Scot
: TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
: PROTEINS, NUCLEIC ACID MOLECULES, AND
: USES THEREOF
: NUMBER OF SEQUENCES: 171
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carol Talkington Verser, Ph.D.
: Heska Corporation
: STREET: 1825 Sharp Point Drive
: CITY: Fort Collins
: STATE: Colorado
: COUNTRY: USA
: ZIP: 80525
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: Wordperfect for Windows, Version 7.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/557,034
: FILING DATE: 21-Apr-2000
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/065,474
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Verser, Carol Talkington
: REGISTRATION NUMBER: 37,459
: REFERENCE/DOCKET NUMBER: HW-5-C1
: TELEPHONE: 970/493-7272
: TELEFAX: 970/484-9505

```

```

? INFORMATION FOR SEQ ID NO: 33:
? SEQUENCE CHARACTERISTICS:
?     LENGTH: 1745 amino acids
?     TYPE: amino acid
?     TOPOLOGY: linear
?     MOLECULE TYPE: protein
?     SEQUENCE DESCRIPTION: SEQ ID NO: 33:
? US-09-557-034-33

```

Query Match	12.9%;	Score 835;	DB 4;	Length 1745;
Best Local Similarity	25.7%;	Pred. No. 4.5e-63;		
Matches 314;	Conservative 154;	Mismatches 452;	Indels 304;	Gaps
QY	34	GSQARALSASSPGGALLAGPGLLLRLALLLAVAAARIMSG---	RC-AGGGACAS	89
DB	35	GESSASFLRAAGNLD-----RVLELL-----	RSQDTINTCANGLNALH	76
QY	90	AAAEAVEPAARELFEACRNGDVERVKRLVTPPEKVNSRDTAGRK-STPLHFAAGFGKDDV	148	
DB	77	ASKEGHHVRELLK--RKADV-----	DAATRKGNLTALHIASLAQGLIV	119
QY	149	EYLLQGANVQAOBDGGLIPLHNACSFGEAEVYVNVLLRHGADPNARDNNWNYTPLHEAAIK	208	
DB	120	TVLVENGANVQSLNGFTPLYMAAQENHESVYRYLLAHNAQALSTEDGFTPLAVALQO	179	
QY	209	GKIDVCIVLLOHCAEPTIRNTDGTALDLA-----DPSAKAVLTGEYKDELLES----	258	
DB	180	GHDVWAVLLENDTRGKVR-----LPALHIAAKKDDTKAATLLQNEHNSDVTSKSGFTPL	235	
QY	259	---ARSGNEKMMALL-----TPLNV-----	NCHAS	281
DB	236	HIAAHYGENVNAOLLEKGANVYQARHNSPLHVATKMGRTNMVSLLLAHGAVIDCRTR	295	
QY	282	DGRKSTPLHLAAGYNRVKIVQLLLOHGADVHAKGDVLPPLHNACSY-----	328	
DB	296	D--LLTPLHCASRSGHDQVVDLLEKGAIPISAKTKNGLAPLHMAAQVDDVDTYLTPLHV	353	
QY	329	---GHYEYVELLVKHGACVNAMDWQFTPLHEAASKNREVCSSLLSYCGADPTLLNCHN	384	
DB	354	AAHGCHVYAKLLLDNRADPNALANGFTPLHITACKNKRIKIVELLKHYA-----	404	
QY	385	KSADLAPTQPKERLAYEFKGH-----SLLQAAAREADVTRIKKHLSEMNWFKHPQTHE	439	
DB	405	--AIEATTESGLSPLHVAAPMGAINIVYLLQGANADVATVR-----	CE	447
QY	440	TALHCAAASPYPRKQICEULLLRKGANINEKYEFTPLPHVASEKAHNDVVEVVKHEAK	499	
DB	448	TPHLHARA---NOTDIVRVLRNGAQVDAARELOTPHLIASRLGNTDITVILLQANAS	504	
QY	500	VNALDNLGOTSLHRAAYCGHLOCTRLLSYGCDPNLIISLQGTALQMGNNVQQLLOEGI	559	
DB	505	PNAATRDLYPLHIAKEQEEVAAALMDHGTDKTLTKRKGFTPLHL-----	551	
QY	560	SLGNSEADQLLEAAKAGDVETVKKLTQVSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY	618	
DB	552	-----AAKGNLPPAKSL--LERTGTPVDIECKNQVTPPLHVAHYNDKVAL	596	
QY	619	LLQHGADVHAKDKGLVPLHNACSYGHEYAEVALLKHGAVNVADLWKETPLHEAAAKG	678	
DB	597	LLENGASAAAAKNGYTPHLHIAAKKNQMDIATSLHYKANANAESKAGFTPLHLAAQEGH	656	
QY	679	YEICKLLLOHGADPTKKNRDGNTPDLVKGDDTDIQDLRGLDAAALLDAAKKGLARVKKL	738	
DB	657	REMAALLIENGAKVGAQRNG-----	677	
QY	739	SSPDNVNCRDQTGRHSTPLHLAAGYNNLEVAEYVLLQHGADVNAQDKGGLIPLHNAASYGH	798	
DB	678	-----LTPMHLCAQEDRVSAEELVYKENAIDPKTKAGYTPPLHVACHFQO	722	
QY	799	VDVAALLIKYNACVNAITDKWAFTPPLHEAAQKGTQLCALLLAHGADPTLLKNGEGOTPLDL	858	
DB	723	INNVRELIHGARSVITRASVYTPHLQAQQQGHNSVYRVLLHGHGASPNVHTSTGTGTPLSI	782	

**QY**   859 VS-----ADVSALLT-----AAMPSPALPCYCYPKOVINGVRSPGATADAL----SSGPSS 90S  
               ||| : |||  
**Dd**   783 AERLGYVSVVEALKTTITVTITTETTYEERYKPQ-----NPEAMNETMFSDESEGE 836  
               ||| : |||  
**QY**   906 PSSLSAASSLNDNLGGSFSELSSVVWSSSG-----TEGASSLEKKKEVPGVDFSIQTQFVNRL 959  
               :||| : |||  
**Dd**   837 DNOITANAHADFESLTk--GLHDSGTGHVLHTATEPILS-ASPEVGSTGDLDLAIrKA 893  
               :||| : |||  
**QY**   960 GLEHL-----MDIFEREQIITDLVLVMHGHELKEIGINAYGH-----RHKLIKGVVER 1006  
               :||| : |||  
**Dd**   894 QHEPIITAMDPDSLDAISLPDNVTI-MRTTMQPQSFLISEFWVDARGCAMGRCH---SGVRI 949  
               :||| : |||  
**QY**   1007 LISGOOQLNP-YILTNTSCGTI-----LIDLSPDKKEFSV----- 1042  
               :||| : |||  
**Dd**   950 IIPPRAPQPTRVTCRYLGKDCLKLAHPPLUSEGEALASRILEMAPHGAKFLPGVILEVPHF 1009  
               :||| : |||  
**QY**   1043 -----EEEMOSTVREHDGGHAGGIENRYNILKIQRVCNKKLWEYTHRRRKEVSEENHNHA 1098  
               :||| : |||  
**Dd**   1010 ASURGREIREIVLRSDDGh---WKehOLEATEDAVQEVLNESFDABELSOLDLITSR 1065  
               :||| : |||  
**QY**   1099 NERMLFHGSFPFNAILHKGFDERH 1122  
               :||| : |||  
**Dd**   1066 ITRILTNDFPMFYFAVTVRVRQEVH 1089  
               :||| : |||

RESULT 11  
US-09-082-059-2  
; Sequence 2, Application US/09082059A  
; Patent No. 6225086  
; GENERAL INFORMATION:  
; APPLICANT: Morriow, Jon S.  
; APPLICANT: Devatarajan, Prasad  
; TITLE OF INVENTION: NO.: 6225086el Ankyrin Proteins and a Method for Their Isolation  
; FILE REFERENCE: 44574-S002-US  
; CURRENT APPLICATION NUMBER: US/09/082\_059A  
; CURRENT FILING DATE: 1998-05-21  
; EARLIER APPLICATION NUMBER: 60/047356  
; EARLIER FILING DATE: 1997-05-21  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: patentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1088  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; OS-09-082-059-2

Query Match	9.7%	Score 629.5;	DB 4;	Length 1088;
Best Local Similarity	28.1%	Pred. No. 1.6e-45;		
Matches 195;	Conservative	85;	Mismatches 234;	
			Indels 181;	Gaps 16;

Qy	304	LLQHGADVHAKDGDVPLHNACSYGHYEVTELLVHKGACVNAMDILWQTPPLHEAASKNR	363
Db	7	LLDKANPNKALNGFTPLPIACKNRKIYMEWELLKHGASIQAVTESGLTPIHVAAFMGH	66
Qy	364	VEVCSLLSYSGADPTLLNCHNKSAIDLAPTQLKERLAYEFKHSLLQAAREADVTRIKK	423
Db	67	VNIYSQLMHGKASPTNVRG-----	87
Qy	424	HLSLEWVFKHPQTHETALHCAASPYPKRKQICELLLRKANINBKTEFTPLHVASE	483
Db	88	-----ETALHMAARS--GQAEVVRVLVDGQAVEAKAKDDQTPPLHISAR	129
Qy	484	KAHNDVVEVVVKHEAKVNALDNLGTSLHRAAYCGHLQTRCLLSYGCPCNIIISLOGETA	543
Db	130	LGRKADIYQOILLQOGASNAATTSSYTPLHLHSAREGHEDVAAFLLDHGASLSITTKKGF--	187
Qy	544	LQMGNENVQOILLQEGISLGNSEADRLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTP	603
Db	188	-----Tp	189

Qy	604	LHFAAGYNRVSVVEYLLQHGADVHARDKGLVPLHNACSYGHVEAELLVKHGAVNVAD	663
	: :		
Db	190	LHVAAYK GKLEVANILLLOKSPDAAGKSGLTPLHVAAYDNOKVALDLLDDOGASPHAAA	249



NAME: Verser, Carol Talkington  
 REGISTRATION NUMBER: 37,459  
 REFERENCE/DOCKET NUMBER: HW-5-C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 970/493-7272  
 TELEFAX: 970/484-9505  
 INFORMATION FOR SEQ ID NO: 139:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 352 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: Protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 139:  
 US-09-557-034-139

Query Match 8.0%; Score 514.5; DB 4; Length 352;  
 Best Local Similarity 23.7%; Pred. No. 2.4e-36;  
 Matches 137; Conservative 64; Mismatches 139; Indels 237; Gaps 4;

QY 131 RKSTPLHFAAGFKDQVVEYLLQNGANVQARDGGGLIPLHNACSFHAEVNVNLLRHGAD 190  
 DB 10 RGETPLHARANQTDIVRVLRNGAQVDAARELQTPHIASELGNTDIVILLQANAS 69  
 QY 191 PNARDNNYTPLEHAAIKGIDVICVILLQHGAEPTINTDGTALDLADPSAKAVLTGEY 250  
 DB 70 PNAATRDLYTPLHAAKEGQEEVAAILMDHCTKTLTKKG----- 110  
 QY 251 KDELLESARSNGNEKMMALLTPLNVNCHASDGRKSTPLHLAGYRVKIVQVLLQHGAD 310  
 DB 111 -----FTPLHAAKYGNLPVAKSLLEGRTP 135  
 QY 311 VHAKDGLVPLHNACSYGHYVETELVKGACVNMADLWQFTPLHEAASKNRVEVCSLL 370  
 DB 136 VDIEGKNQVTPLVAAHYNDKVALLLLENGASAHAAKNGYTPPLHIAAKNQMDIASTL 195  
 QY 371 LSYGADPTLLNCHNKSALDAPTQPKERLAYEPKGHSLLOAAAREADVTRIKHLSLEMY 430  
 DB 196 LHY----- 198  
 QY 431 NFXHPQTHETALHCAASPYKPKQICEALLRKGANINEKTKELTPLHVASEKAHNDVV 490  
 DB 199 -----KANAAESKAGFTPLHLAAQEGHREMA 225  
 QY 491 EVVYKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIISLOGFTALQMGNE 550  
 DB 226 ALLIENGAKVGA----- 237  
 QY 551 VQQLQEGISLGNSEADRLLEAAKAGDVTYKVKLCTVQSVNCRDIEGROSTPLHFAAGY 610  
 DB 238 -----QARNG-----LTPMHLCAQE 252  
 QY 611 NRVSVEYLLQHGADVHAKDKGGLVPLHNACSYGHYVEAEVLLYKHGAVNVADLWKFTPL 670  
 DB 253 DRVSVAELVKNENAIADPKTKAGYTPLVACHFGQINWVFLIEHGARSVITRASYTPL 312  
 QY 671 HERAAKGYEICKLLQHGADPTKKNRDGNTPDLVK 707  
 DB 313 HQAAQOQHNSVRYLLEHGASPNVHTSTGTPTLSIAE 349

RESULT 14  
 US-09-031-485-23  
 ; Sequence 23, Application US/09031485  
 ; Patent No. 5824306  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Liang  
 ; APPLICANT: Blehm, E. Scot  
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
 ; TITLE OF INVENTION: USES THEREOF  
 ; NUMBER OF SEQUENCES: 85  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.

ADDRESSEE: Heska Corporation  
 STREET: 1825 Sharp Point Drive  
 CITY: Fort Collins  
 STATE: Colorado  
 COUNTRY: USA  
 ZIP: 80525  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: Windows 95  
 SOFTWARE: WordPerfect for Windows, Version 7.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/031,485  
 FILING DATE:  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/847,429  
 FILING DATE: 24-APR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Verser, Carol Talkington  
 REGISTRATION NUMBER: 37,459  
 REFERENCE/DOCKET NUMBER: HW-5  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 970/493-7272  
 TELEFAX: 970/484-9505  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 303 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-031-485-23

Query Match 6.5%; Score 421.5; DB 2; Length 303;  
 Best Local Similarity 30.4%; Pred. No. 2.3e-28;  
 Matches 119; Conservative 52; Mismatches 125; Indels 95; Gaps 8;

QY 467 INEKTKEFLTPHVASEKAHNDVVEVYVVKHAKVNALDNLGQTSLHRAAYCGHLQTCRL 526  
 DB 1 VDDVTVDYLTPLHV-----AAHCGHVRVAKLL 27  
 QY 527 LSYGCDPNIIISLOGFTALQMGNEVQQLQEGISLGNSEADRLLEAAKAGDVTYKVKLC 586  
 DB 28 LDRNADPNARALNGFTPLHI-----ACKNRKIKIVELL 61  
 QY 587 TVQSVNCRDIEGRQS---TPLHEAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 643  
 DB 62 KYHAA---IEATTESGLSPLHVAFAFMCAINIVYLLQOGANADVATYRGETPLHLAARA 117  
 QY 644 GHYEAEVLLYKHGAVNVADLWKFTPLHEAAKAGKYEICKLLLOHGADPTKKNRDGNTP 703  
 DB 118 NOTDIVRVLRNGAQVDAARELQTPHIASELGNTDIVILLQANASPNAAATRDLYTPL 177  
 QY 704 DL-VKDGDTDIDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTCGRHSTPLHL 759  
 DB 178 HIAAKEGQEEVAAILMDHGTDTLL--TKKG-----FTPLHL 212  
 QY 760 AAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAAALLIKYNACVNATDKWA 819  
 DB 213 AAKYGNLPVAKSLLEGRTPVDIEGKNQVTPLVHAAHYNDKVALLLLENGASAHAAKNG 272  
 QY 820 FTPLHEAAKAGRTQICALLLHAGADPTLKNG 850  
 DB 273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303

RESULT 15  
 US-08-847-429A-23  
 ; Sequence 23, Application US/08847429A  
 ; Patent No. 5827692  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Liang  
 ; APPLICANT: Blehm, E. Scot

Search completed: February 12, 2003, 12:27:27  
Job time : 25.9573 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: February 12, 2003, 03:33:09 ; Search time 25.4359 Seconds  
(without alignments)  
1245.508 Million cell updates/sec

Title: US-09-843-159B-4  
Perfect score: 6464  
Sequence: 1 RCSARRGAAGGCGAQRGARV.....AYPEYLITYQIMRPGMVDG 1240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5602	86.7	1074	10	US-09-509-196A-2
2	5103.5	79.0	1327	10	US-09-841-835-2
3	3938	60.9	802	9	US-09-964-899-41
4	3307	51.2	949	10	US-09-841-835-10
5	2132.5	33.0	673	10	US-09-841-835-8
6	815.5	12.6	1724	9	US-09-964-899-43
7	401	6.2	426	10	US-09-908-711-70
8	398.5	6.2	1719	9	US-10-012-896-378
9	398.5	6.2	1719	9	US-09-895-793-378
10	398.5	6.2	1719	9	US-09-895-814-378
11	398.5	6.2	1719	10	US-09-759-143-378
12	398.5	6.2	1719	10	US-09-780-669-378
13	398.5	6.2	1719	10	US-09-822-827-378
14	392	6.1	740	10	US-09-835-788A-12
15	332.5	5.1	835	10	US-09-947-199-2
16	324.5	5.0	835	10	US-09-947-199-8
17	312.5	4.8	599	10	US-09-735-368-2
18	297	4.6	551	10	US-09-835-788A-17
19	281.5	4.4	285	10	US-09-835-788A-18

20	278.5	4.3	1054	10	US-09-798-042-87	Sequence 87, Appl
21	268	4.1	306	10	US-09-835-788A-19	Sequence 19, Appl
22	258.5	4.0	323	10	US-09-884-870-2	Sequence 2, Appli
23	258	4.0	251	10	US-09-835-788A-13	Sequence 13, Appl
24	247.5	3.8	328	10	US-09-758-593A-11	Sequence 11, Appl
25	247.5	3.8	328	10	US-09-758-593A-12	Sequence 12, Appl
26	247.5	3.8	329	10	US-09-880-192-62	Sequence 62, Appl
27	247.5	3.8	329	10	US-09-758-593A-1	Sequence 1, Appli
28	243.5	3.8	2444	10	US-09-944-849-2	Sequence 2, Appli
29	243	3.8	752	10	US-09-927-180-2	Sequence 2, Appli
30	240	3.7	671	9	US-09-924-400-306	Sequence 306, Appl
31	240	3.7	671	9	US-10-012-896-380	Sequence 380, App
32	240	3.7	671	9	US-09-895-793-380	Sequence 380, App
33	240	3.7	671	9	US-09-895-814-380	Sequence 380, App
34	240	3.7	671	10	US-09-825-301-10	Sequence 10, Appl
35	240	3.7	671	10	US-09-759-143-380	Sequence 380, App
36	240	3.7	671	10	US-09-780-669-380	Sequence 380, App
37	240	3.7	671	10	US-09-810-936-306	Sequence 306, App
38	240	3.7	671	10	US-09-822-827-380	Sequence 380, App
39	240	3.7	671	10	US-09-429-755-306	Sequence 306, App
40	237.5	3.7	656	9	US-09-924-400-305	Sequence 305, App
41	237.5	3.7	656	9	US-10-012-896-379	Sequence 379, App
42	237.5	3.7	656	9	US-09-895-793-379	Sequence 379, App
43	237.5	3.7	656	9	US-09-895-814-379	Sequence 379, App
44	237.5	3.7	656	10	US-09-825-301-9	Sequence 9, Appli
45	237.5	3.7	656	10	US-09-759-143-379	Sequence 379, App

ALIGNMENTS

RESULT 1

US-09-509-196A-2

; Sequence 2, Application US/09509196A

; Patent No. US20020037582A1

; GENERAL INFORMATION:

; APPLICANT: DALY, Roger J.

; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling

; TITLE OF INVENTION: Proteins

; FILE REFERENCE: 1871-129

; CURRENT APPLICATION NUMBER: US/09/509,196A

; PRIOR FILING DATE: 2000-03-23

; PRIOR APPLICATION NUMBER: P09388

; PRIOR FILING DATE: 1997-09-23

; PRIOR APPLICATION NUMBER: PCT AU98/00795

; PRIOR FILING DATE: 1998-09-23

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1074

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-509-196A-2

Query Match 86.7%; Score 5602; DB 10; Length 1074;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1068; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy	167	IPLHNACSFHAEVYVNLRLRHGADPNARDNNWYTPLHEAAIKGKIDVCTIVLQHGAEPTI	226
Db	1	IPLHNACSFHAEVYVNLRLRHGADPNARDNNWYTPLHEAAIKGKIDVCTIVLQHGAEPTI	60
Qy	227	RNTDGR TALDPSAKAVLTGEYKKDELLESARGNEKMMALLTPLNVNCHASDGRKS	286
Db	61	RNTDGR TALDPSAKAVLTGEYKKDELLESARGNEKMMALLTPLNVNCHASDGRKS	120
Qy	287	TPHLHAACYNRVKIVOLLQHCADVHAKDGLVPLHNACSYGHYEVTLLVKGACVNA	346
Db	121	TPHLHAACYNRVKIVOLLQHCADVHAKDGLVPLHNACSYGHYEVTLLVKGACVNA	180
Qy	347	MDLWQFTPLHEAASKNRVEVCSLLSYGADPTLLCHNKSAITDLAPTOLKRLAYEFGK	406

Db 181 MDLWQFTPLHGAASKNRVEVCSLLLSYGADPTLLNCKNKAIDLAPTPQLKERLAYEFGK 240  
QY 407 HSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAAASYPKPKQICELLKRGAN 466  
Db 241 HSLQAAAREADVTRIKKHLSELMVNFKHPQTHETALHCAAAASYPKPKQICELLKRGAN 300  
QY 467 INEKTKEFLTPLHVASKAHNDVVVVVVKHEAKVNDLNGQSLHRAAYCGHLQTCRL 526  
Db 301 INEKTKEFLTPLHVASKAHNDVVVVVVKHEAKVNDLNGQSLHRAAYCGHLQTCRL 360  
QY 527 LSYGCDPNIISLOGFTALOMGNVVOQLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 586  
Db 361 LSYGCDPNIISLOGFTALOMGNVVOQLQEGISLGNSEADROLLLEAAKAGDVETVKKLC 420  
QY 587 TVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 646  
Db 421 TVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480  
QY 647 EVAELLVKHGAVNVADLWKFTPLHFAAAKGYEICKLLOHGADPTKKNRDGNTPDLV 706  
Db 481 EVAELLVKHGAVNVADLWKFTPLHFAAAKGYEICKLLOHGADPTKKNRDGNTPDLV 540  
QY 707 KDGDTDIDLGRDAALLDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNL 766  
Db 541 KDGDTDIDLGRDAALLDAAKKGCLARVKLSSPDNVNCRDTQGRHSTPLHLAAGYNL 600  
QY 767 EVAELLVKHGADYNAQDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWAFPLHEA 826  
Db 601 EVAELLVKHGADYNAQDKGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFPLHEA 660  
QY 827 AQKGRQLCALLLAHGADPTLKNQEGQTPLDIVSADDSALLTAAMPSPALPSCYKQV 886  
Db 661 AQKGRQLCALLLAHGADPTLKNQEGQTPLDIVSADDSALLTAAMPSPALPSCYKQV 720  
QY 887 NGVRSFCATADALSSGSPSSLSAASSLDNLGSFSELSSVSSVSSGTEGASSLEKKEVP 946  
Db 721 NGVRSFCATADALSSGSPSSLSAASSLDNLGSFSELSSVSSVSSGTEGASSLEKKEVP 780  
QY 947 GVDFTSQFVRNLGLEHMLDIFEREQITLDVLEMGHKKELKEIGINAYGHRHKLKIGVER 1006  
Db 781 GVDFTSQFVRNLGLEHMLDIFEREQITLDVLEMGHKKELKEIGINAYGHRHKLKIGVER 840  
QY 1007 LISGQOGLNPLYLNTSGSGTILIDLSPPDKFQSVSEEMQSVTRHRDGGHAGGIFNRY 1066  
Db 841 LISGQOGLNPLYLNTSGSGTILIDLSPPDKFQSVSEEMQSVTRHRDGGHAGGIFNRY 900  
QY 1067 NILKIOKVCNKKLWERYTHRRKEVSEENHNANERMLFHGSPFVNAILHKGFDERHAYIG 1126  
Db 901 NILKIOKVCNKKLWERYTHRRKEVSEENHNANERMLFHGSPFVNAILHKGFDERHAYIG 960  
QY 1127 GMFGAGIYFAENSCKSNQYVYIGGGTGPCVHKDRSCYICHRQLLCFVRTVLGKSFQFSA 1186  
Db 961 GMFGAGIYFAENSCKSNQYVYIGGGTGPCVHKDRSCYICHRQLLCFVRTVLGKSFQFSA 1020  
QY 1187 MKMAHSPGPHSHSVTGRPSVNGALAEYVYIRGQAPPEYLITYQIMRPGMWDC 1240  
Db 1021 MKMAHSPGPHSHSVTGRPSVNGALAEYVYIRGQAPPEYLITYQIMRPGMWDC 1074

RESULT 2

US-09-841-835-2  
; Sequence 2, Application US/09841835  
; Patent No. US20020076795A1

; GENERAL INFORMATION:

; APPLICANT: de Lange, Titia

; APPLICANT: Smith, Susan

; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRP1 AND METHODS

; TITLE OF INVENTION: OF USE THEREOF

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1327 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOPHETICAL: NO

Query Match

Best Local Similarity 79.0%; Score 5103.5; DB 10; Length 1327;  
Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAHTGAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRCA 81  
Db 112 SAAGVAPNPAGSGNNSPSSSSPTSSSSSSPSGPG-----SSLAESPAAAGVSSSTAPL 165  
QY 82 GGAACAASAAAEAVEPAARELFEACRNGDVERVKRLVTPPEKVNRSRTAGRKSTPLHPAAG 141  
Db 166 GPGAGPGTGVPAVSGALRELLEACRNGDVSRYKLVDAANVANAKDMAGRKSSPLHPAAG 225  
QY 142 FGRKDVVEYLLQNGANVOARDGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTP 201  
Db 226 FGRKDVVEYLLQNGANVHARDGGLIPLHNACSFHAEVNVNLLRCCQGADPNARDNNWYTP 285  
QY 202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESARS 261  
Db 286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGTALDADPSAKAVLTGEYKKDELLESARS 345  
QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVLP 321  
Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVLP 405  
QY 322 LHNACSYGHYEVTELLVKGACVNMADLWQFTPLHAAASKNRVEVCSSLLSYGADPTLLN 381  
Db 406 LHNACSYGHYEVTELLVKGACVNMADLWQFTPLHAAASKNRVEVCSSLLSYGADPTLLN 465  
QY 382 CHNKSIDLAPTPOLKERLAYEFKHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHETA 441  
Db 466 CHKSAVDMNAPTPELRELTTEYEFKHSLLQAAAREADVTRIKKHLSELMVNFKHPQTHETA 525  
QY 442 LHCAASAPYKPKQICELLKRGANINEKTEKFTPLPLHVASEKAHNDVVEVVKHAKVN 501  
Db 526 LHCAVASLHPKQVQVTELLLRKGANINEKTEKFTPLPLHVASEKAHNDVVEVVKHAKVN 585  
QY 502 ALDNLGOTSILHRAAYCGHLQTCRLLSYGCDPNIIISLQGTALQMGNNVVOQLQEGISL 561  
Db 586 ALDNLGOTSILHRAAYCGHLQTCRLLSYGCDPNIIISLQGTALQMGNNVVOQLQEGISL 645

QY 562 GNSEADRLLEAAKAGDVETVKKLTQVSNCRDIEGRQSTPLHFAAGYNNRVSVVEYLQ 621  
Db 646 RTSDVYRLLEAKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNNRVSVVEYLH 705  
QY 622 HCADVHAKDKGLVPLHNACSYGHEVAELLYKHGAVNVADLWKFPTPLHFAAAGKGYEI 681  
Db 706 HCADVHAKDKGLVPLHNACSYGHEVAELLYKHGASVNVADLWKFPTPLHFAAAGKGYEI 765  
QY 682 CKLLHQHGADPTKKNRDNPTLDLVKDGDTDIDQLLRLGGDAALLDAACKGCLARVKKLSPP 741  
Db 766 CKLLKHGADPTKKNRDNPTLDLVKDGDTDIDQLLRLGGDAALLDAACKGCLARVKKLTCTP 825  
QY 742 DNVCNCRDTQGRHSTPLHFAAGYNNLEVAEYLLQHGADVNAQDKGLPLHNAASYCHVDV 801  
Db 826 ENINCRTQGRNSTPLHFAAGYNNLEVAEYLLQHGADVNAQDKGLPLHNAASYCHVDV 885  
QY 802 AALLIKYNACVNATDKWFTPLHFAAOKGRTOLCALLLAHGADPTLKNOEGOTPLDLVSA 861  
Db 886 AALLIKYNACVNATDKWFTPLHFAAOKGRTOLCALLLAHGADPTLKNOEGOTPLDLVSA 945  
QY 862 DVVSALLTAAMPSPALPSCYKQVNLGVKSPGATADALSSGSPSSLSAASSLNLGSGS 921  
Db 946 DDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTPSCLSAASSIDNLGTP 998  
QY 922 FSLSVSVSSSTEGASLEKK--EVPQVDFSIITQFVRNLGLEHMDIFEREQITLDVLV 979  
Db 999 LAELAVGASNAAGDGAAGTERKEGEVAGLDNMISQFLKSLGLEHLRDIPTETQITLDVLA 1058  
QY 980 EMGHKEKEIGINAYGHRKHLKGVERTLSGOOGLNPYLTNTSGTILIDLPDDKEF 1039  
Db 1059 DMGHKEKEIGINAYGHRKHLKGVERTLSGOOGLNPYLTNTSGTILIDLPDEKEY 1118  
QY 1040 OSVEEMQSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEENHNAN 1099  
Db 1119 OSVEEMQSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRKEVSEENHNAN 1178  
QY 1100 ERMPLHGSFPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHK 1159  
Db 1179 ERMPLHGSFPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVYIGGGTGCPCVHK 1238  
QY 1160 DRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHSVTRGPRSVNGLAELAEYVYIRGE 1219  
Db 1239 DRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPGHSVTRGPRSVNGLAELAEYVYIRGE 1298  
QY 1220 QAYPEYLITYQIMRPE 1235  
Db 1299 QAYPEYLITYQIMRPE 1314

RESULT 3  
US-09-964-899-41  
; Sequence 41, Application US/09964899  
; Patent No. US2002017446A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Dalia et al.  
; TITLE OF INVENTION: Identification of Genes Involved in  
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster  
; FILE REFERENCE: 4-31612 A  
; CURRENT APPLICATION NUMBER: US/09/964,899  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/236,893  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/298,309  
; PRIOR FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41  
; LENGTH: 802  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-964-899-41

Query Match 60.9%; Score 3938; DB 9; Length 802;

Best Local Similarity 67.2%; Pred. No. 1.9e-270;  
Matches 787; Conservative 4; Mismatches 5; Indels 376; Gaps 5;  
QY 75 MSGRRCAGGAGAACASAAAFAVEPAARELFACRNGDVERVKRLVTPKEVNSRDTAGRKST 134  
Db 1 MSGRRCAGGAGAACASAAAFAVEPAARELFACRNGDVERVKRLVTPKEVNSRDTAGRKST 60  
QY 135 PLHFAAGFRKDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVYVNLRLRHGADPNAR 194  
Db 61 PLHFAAGFRKDVVEYLLQNGANVQARDGGLIPLHNACSFHAEVYVNLRLRHGADPNAR 120  
QY 195 DNWNTPLHEAATKGDIVCIVLLOHGABPTIRNTDGRALTDLADPSAKAVLTGE----- 249  
Db 121 DNWNTPLHEAATKGDIVCIVLLOHGABPTIRNTDGRALTDLADPSAKAVLTGKSVYS 190  
QY 250 YKKDELLES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNNRVKIVOLLQHG 308  
Db 181 YSRKPKVKNLARSNEEKMMALLTPLNVNCHASDGR----- 216  
QY 309 ADVHAKDKGLVPLHNACSYGHEVTELLVKHGACVNAAMDLMQFTPLHFAASKNRVEVCS 368  
Db 217 -----KHGACVNAAMDLMQFTPLHFAASKNRVEVCS 246  
QY 369 LLSYGADPTLLNCHNKSALDAPTPQLKERLAYEFKHSLLQAAAREADVTRIKKHLISLE 428  
Db 247 LLSYGADPTLLNCHNKSALDAPTPQLKERLAYEFKHSLLQAAAREADVTRIKKHLISLE 306  
QY 429 MYNFKHQPTHETALHCAASAPYKPKKOICELELLRKGANINEKTEKPLTPLHVAASEKAHND 488  
Db 307 MYNFKHQPTHETAL----- 320  
QY 489 VVEVVVKHEAKNALDNLGQTSILHRAAYCGHLQTCRLLLSYGCDPNIISLQGTALQMGN 548  
Db 321 ----- 320  
QY 549 ENVQQLLOBEGISLGNSEADRLLEAAKAGDVETVKKLTQVSNCRDIEGRQSTPLHFAA 608  
Db 321 -----KLCTVQVSNCRDIEGRQSTPLHFAA 345  
QY 609 GYNNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEVAELLYKHGAVNVADLWKF 668  
Db 346 GYNNRVSVVEYLLQHGADVHAKDKGLVPLHNACSYGHEVAELLYKHGAVNVADLWKF 405  
QY 669 PLHFAAAGKGYEICKLLOHGADPTKKNRDNPTLDLVKDGDTDIDQLLRLGGDAALLDAK 728  
Db 406 PLHFAAAGKGYEICKLLOHGADPTKKNRDNPTLDLVKDGDTDIDQLLRLGGDAALLDAK 465  
QY 729 KGCLARVKKLSPPDNVNCNCRDTQGRHSTPLHFAAGYNNLEVAEYLLQHGADVNAQDKGLI 788  
Db 466 KGCLARVKKLSPPDNVNCNCRDTQGRHSTPLHFAAGYNNLEVAEYLLQHGADVNAQDKGLI 525  
QY 789 PLHNAASYCHVDVVAALLIKYNACVNATDKWFTPLHFAAOKGRTOLCALLLAHGADPTLK 848  
Db 526 PLHNAASYG----- 534  
QY 849 NOEGOTPLDLVSADDDVALLTAAMPSPALPSCYKQVNLGVKSPGATADALSSGSPSS 908  
Db 535 ----- 534  
QY 909 LSAASSLNLGSGFSLSVSSSSTEGASLEKKEVPGVDFSIITQFVRNLGLEHMDIF 968  
Db 535 ----- 534  
QY 969 EREQITLDLVEMGHKELKEIGINAYGHRKHLKGVERTLSGOOGLNPYLTNTSGSTI 1028  
Db 535 -----ITLDLVEMGHKELKEIGINAYGHRKHLKGVERTLSGOOGLNPYLTNTSGSTI 590  
QY 1029 LIDLSPDDKEFOSVEEMQSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRK 1088  
Db 591 LIDLSPDDKEFOSVEEMQSTVREHRDGGHAGGIFNRYNLTQKVCNKKLWERYTHRRK 650  
QY 1089 EYSEENHNANERMLFHGSPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVY 1148  
Db 1089 EYSEENHNANERMLFHGSPFVNAIITHKGFDERHAYIGMFGAGIYFAENSSKSNQYVY 1148

Db 651 EVSEENHNANERMLPHGSPFVNATIIHKGFDERHAYIGMGFGAGIYFAENSCKSNQYVYG 710  
QY 1149 IGGTCCPVHKDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPPGHSHVTCGRPSVNGL 1208  
Db 711 IGGTCCPVHKDRSCYICHRQLLFCRVTLGKSLQFSAMKMAHSPPGHSHVTCGRPSVNGL 770  
QY 1209 ALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240  
Db 771 ALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 802

RESULT 4  
US-09-841-835-10  
; Sequence 10, Application US/09841835  
; Patent No. US20020076795A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 949 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-841-835-10

Query Match 51.2%; Score 3307; DB 10; Length 949;  
Best Local Similarity 76.2%; Pred. No. 1e-225;  
Matches 638; Conservative 72; Mismatches 119; Indels 8; Gaps 2;  
QY 22 AAGTADPDVYTAGQAARALSSPGGLALLAGPGLLLRLALLLAAVAAARIMSGRCA 81  
Db 112 SAAGVAPNAGSGNNSSPSSTSSSSPSPG-----SSLAESPEAGVSTAPL 165  
QY 82 GGAACASAAAEAVEPAARELFACRNGDVERVKRLVTPKVNSTRDTAGRKSTPLHFAAG 141  
Db 166 GPGAAGGTGVPAVAGRLRELLACRNGDVRVKRLVDAANVNAKMDAGKRSSPLHFAAG 225  
QY 142 FGRKDVVYLLQNGANYQARDGGLIPLHNACSGHAEVYNNLLRHGADPNARDNNWYTP 201  
Db 226 FGRKDVVYLLQNGANYQARDGGLIPLHNACSGHAEVYNNLLRHGADPNARDNNWYTP 285  
QY 202 LHEAAIKGKIDVICIVLQHGADPNIIRNTDGTALDLADPSAKAVLTGEYKKDELLESARS 261

Db 286 LHEAAIKGKIDVICIVLQHGADPNIIRNTDGTALDLADPSAKAVLTGEYKKDELLESARS 345  
QY 262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVP 321  
Db 346 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVOLLQHGADVHAKDKGLVP 405  
QY 322 LHNACSYGHYEVTELLVKHGCAGVNDMLQWFTPLHFAASKNRVVEVCSLLLSYGADPTLN 381  
Db 406 LHNACSYGHYEVTELLVKHGCAGVNDMLQWFTPLHFAASKNRVVEVCSLLLSHGADPTLN 465  
QY 382 CHNKSALDAPTPOLKERLAYEFKSHLSLOAAREADVTRIKKHLSEMVNFKHPQTHETA 441  
Db 466 CHGKSAVDMAPTDELRLTYEFKSHLSLOAAREADLAKYKKTALETIIRFKQPQSHETA 525  
QY 442 LHCAASAPYKPKKOICELLRLKGANINEKTEFLTPLVASEKAHNDVVEVYVYKHAKVN 501  
Db 526 LHCASVSLHPKPKQVTELLRLKGANYNKNDPMTPLHVAERAHNDVMEVLKHGAKMN 585  
QY 502 ALDNLGQTSIHLAAYCCHLQTCRLLSYSGDPNIIISLQGTALOMGNENYQQLQEGISL 561  
Db 586 ALDNLGQTSIHLAAYCCHLQTCRLLSYSGDPNIIISLQGTALOMGNENYQQLQEGISL 645  
QY 562 GNSEADROLLEAAKAGDVETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVYELLQ 621  
Db 646 RTSVDVYRLLEASKAGDLETVKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVYELLQ 705  
QY 622 HGADVHAKDKGLVPPLHNACSYGHYEAELLVKHGVNVDLWKFTPLHFAAAKGYEI 681  
Db 706 HGADVHAKDKGLVPPLHNACSYGHYEAELLVKHGVNVDLWKFTPLHFAAAKGYEI 765  
QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDLRLGDAALLDAKKGCLARVKLSSP 741  
Db 766 CKLLKHGADPTKKNRDGNTPLDLVKDGTDIQDLRLGDAALLDAKKGCLARVKLCTP 825  
QY 742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799  
Db 826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGCLA 885  
QY 800 DVAALLIKYNACVNATDKWFTPLHFAAAKGRTOCLALLAHAGADPTLKNOEGQTPL 856  
Db 886 RVOKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPL 942

RESULT 5  
US-09-841-835-8  
; Sequence 8, Application US/09841835  
; Patent No. US20020076795A1  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/841,835  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,387  
; FILING DATE:





Db 78 LHLASKGHHVEVSELLQREANVDAAATKGN TALHIA SLAQAEVVKVLTNGANVNAQS 137  
QY 196 NWNVTPLHEAAIKGIDVCI VLLQHGAEPTIRNTDGR TALDLPDPSAKAVLTGEYKKDEL 255  
Db 138 QNGFTPLYMAXQENHLEWVFLDNGASQXLXTEDGTPLAVA----- 180  
QY 256 LESARSENEEKMMALLTPLNVNCHASGRKSTP-LHLAAGYNNRVKIVOLLQHGADVHAK 314  
Db 181 ---LQOQHDQVVSLLLE-----NDTKGKVRPLPALHTAARKDDTKAAALLIQDNNDNADVE 231  
QY 315 DKGLDPLVHNACSYGHYEVTELLVKKHCACYNAMDLPQFTPLHEAASKNRVEVCSELLLSYG 374  
Db 232 SKSGFTPLHTAAHGNINATLNLNRXAAVDFTARNDDITPLHVASKRGNAWMVKLLDRG 291  
QY 375 ADPTLLNCHNKSADLAPTQPKERLAYEFKHSLLQAAAREADVTRIKKHSLEWVNFKH 434  
Db 292 A-----KIDAKTROGL----- 302  
QY 435 POTHETALHCAASPYPKRQICELLRKGANINEKTEFTPLHVASEKAHNDVVEVV 494  
Db 303 ----TPLHCGARGSH---EQVVEMLLDRAAPILSKTKNGLSPLHMATQGDHLNCVOLL 354  
QY 495 KHEAKVNALDNGOTSLHRAAYCGHLOTCRLLL 527  
Db 355 OHNVPPVDVTDNYLTALHVAHCHGYKAVKVL 387

## RESULT 8

US-10-012-896-378

; Sequence 378, Application US/10012896

; Publication No. US20020183251A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillion, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Meagher, Madeleine Joy

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C27

; CURRENT APPLICATION NUMBER: US/10/012,896

; CURRENT FILING DATE: 2001-12-10

; NUMBER OF SEQ ID NOS: 1011

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 378

; LENGTH: 1719

; TYPE: PR

; ORGANISM: Homo sapiens

US-10-012-896-378

Query Match

Best Local Similarity 6.2%; Score 398.5; DB 9; Length 1719;

Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;

QY 80 CAGGAACASAAAEVAPAREL-----FEACRNGDVERV-----KRLVTP 120  
Db 72 CRGSGKSNVSGASGDHDDSAMKTLRNKMGKWCCHCPCPCRGSGSKVGAWDYDSDAFMEP 131  
QY 121 EKVNSRDTAGKSTPLHFAAGFG---RKDVVEYLQNGANYOARDGGLIPLHNAACSGH 177  
Db 132 R-----YHVRGDDLLKHRAAWMGKVPKRDLI--VMLRDTDVNKKDKRKTALHLASANG 185  
QY 178 AEVNNLLRHGADPNARDNNWYTPLEHAAIKGIDVCI VLLQHGAEPTIRNTDGR TALDL 237  
Db 186 SEVVKLLDRRCQLNVLDNKKRTALIKAVQCEDEBCALMLLEHGTDNPIDEYGNWTTLHY 245  
QY 238 ADPSAKAVLTGEYKKDELLESARSNEEKMMALLTPLNVNCHASGRKSTPLHLAAGYNNR 297  
Db 246 A-----TYNEDKLMA----- 255  
QY 298 VKTVOLLQHGADVHAKDKGLDPLVHNACSYGHYEVTELLVKKHCACYNAMDLPQFTPLHE 357  
Db 256 ----KALLYGADIESKNNKHGLTPLLGVHEQKQOVVKFLIKKANLANALDRYGR TALIL 311  
QY 358 AASKNRVEVCSELLLSYGADPTLLNCHNKSADLAPTQPKERLAYEFKHSLLQAAARE 415  
Db 312 AVCCGSASIVSLLLEQNTDVSQDLSGQTAREYAVSSHVHCICQLLSYKEKQMLKISSE 371  
QY 416 ----ADVTRIKKHLSLEMVNFKHPQTH---ETALHCAAS---PYPKRQICELLRLRKA 465  
Db 372 NSNPENVSRT-----NKPRTHMVVEVDSMPAASSVKKPPGLRSKMGKWCRCRCP 421  
QY 466 NINEKTEFTPLHVASEKAHNDVVEVVVHKEAKVNALDNGOTSLHRAAYCGHLOTCRL 525  
Db 422 CCRESGKS---NVGTSGDHDD-----SAMKTLRSKMGK-----WCRHCFPC-- 459  
QY 526 LLSYCCDPNIISLOGFTALQMNENVOQLLOEGISLGNSEADROLLEAAKAGDVETVKL 585  
Db 460 ----CRGS-----GKSNV-----GASGDHDDSAMKTLR-----NKMKGW 489  
QY 586 CTQVSVCNR-----DIEGRSTPLHFAAGYNNRVSVVEYLLQ- 621  
Db 490 CCHCFPCRGSGSKVGAWDYDSDAFMEPRYHVRGDDLDKLHRAAWMGKVPKRDLI VYL 549  
QY 622 HGADVHAKDKGLVPLHNACSYGHYEVAELLVKHGAVNVADLWKFTPLHRAAAKGYEI 681  
Db 550 RDTDVNKKDKOKRTALHLASANGSEVVKLLDRRCQLNVLDNKKRTALIKAVQCEDEC 609  
QY 682 CKLLLOHCADPTKKNRQNTPLDLVKDGTDTIQLLRGDAALLDAAKKGCLARVKKLSSP 741  
Db 610 ALMLLEHGTDNPIPEYGN----- 628  
QY 742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADYNAODKGLIPLHNAASYGHVDV 801  
Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGVHEQKQOV 675  
QY 802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTOLCALLAHGADPTLKNQEGOTPLDLVSA 861  
Db 676 VKFLIKKANLALDRYGR TALILAVCCGSASIVSLLLEQNTDVSQDLSGQTAREYAVS 735  
QY 862 DDVSALLTAAPPSPALPSCYKPOVLNGVYRSPGATADALSSGSPSSLSAASLDNLGSG 921  
Db 736 SHHHV-----COLJSDYKEKQMLKISSE-----NSNPEQDLKLTSEESORFSGS 781  
QY 922 FSELSSVVSSSGTEGASLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLVDVLEM 981  
Db 782 -----ENSQPEKMSQPEINKD-----GPRE-----VE- 804  
QY 982 GHKELKEIGINAYGHRHKLIGVERLISGQOQLNPYLTNTSGSGTILIDLSPDDKEFQS 1041  
Db 805 --EEMKKHESNNVGLLENLTNGV-TAGNCDNGLIPORKSRT-----PENQOFPD 850  
QY 1042 VEEEMQSTVREHROGGHAGGIFNRYNILKIOKVCNK-----KLWERYTHRRKEVSEE 1093  
Db 851 NESEYHRICE-----LVSDYKEKQMPKYSSENPEQDLKLTSEESORLEGSEN 901  
QY 1094 NNNH-----ANERMLFHGSPFVNAILHKGFE 1120

Db 902 GOPELENFMAIEEMKKHGS-----THVGFPPE 927

RESULT 9  
US-09-895-793-378

; Sequence 378, Application US/09895793  
; Publication No. US20020192763A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals de Bassols, Carlota  
; APPLICANT: Foy, Teresa  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C2  
; CURRENT APPLICATION NUMBER: US/09/895,793  
; CURRENT FILING DATE: 2001-06-29  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 378  
; LENGTH: 1719  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-895-793-378

Query Match 6.2%; Score 398.5; DB 9; Length 1719;  
Best Local Similarity 21.3%; Pred. No. 1.4e-19;  
Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;

QY 80 CAGGGACASAAAEVPEAREL-----FEACRNGDVERV-----KRLVTP 120  
Db 72 CRGSGKSNVSGAGDHDSDSANKTLRNMGKWCCHCFCCRGSGKSKYGAWGDDYDSDAFMEP 131  
QY 121 EKVNSRDTAGRKSTPLHFAAGFG---RKDVVEYLLQNGANVQARDGGLIPLHNACSFHG 177  
Db 132 R----YHVRGEDLDKLHRAAWGKVPKDLI--VMLRDTDVNKKQKQKRTALHLASANGN 185  
QY 178 AEVNNLLRHGADPNARNWNYPPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRATDLD 237  
Db 186 SEVVKLLDRRCOLNVDNKKRTALIKAVQCQDECALMLEHGTDPNIPDEYGNTHLY 245  
QY 238 ADPSAKAVLTGEYKDKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNR 297  
Db 246 A-----IYNEDKLMA----- 255  
QY 298 VKTVOLLQHGADVHAKDGLVPLHNACSYGHYEYTELLVKGACVNMADLWQFTPLHE 357  
Db 256 ----KALLYGADIESKNHGLTPLLGVHEQKQVVKELIKKANLNALDRYGRATALIL 311  
QY 358 ASKNRVECSLLLSYGADPTLLNCHINKSAIDLAPTP--OLKERLAYEPKGHSLLOAARE 415  
Db 312 AVCCGSASIVSLLLEONIDVSSODLSGQAREYAVSSHVHCVCQLLSDYKEKQMLKISSE 371  
QY 416 ----ADVTRIKKHLSEMYNFKHPQTH---ETALHCAAS---PYPKRQKQICELLRLKGA 465

Db 372 NSNPENVSRT-----NKPRTHMVVEVDSMAAASVKKPFGGLRSMKMGWCRCFP 421  
QY 466 NINEKTEFLTPLHVASEKAHNDVVEVVKHEAKVNALDNLGOTSLHRAAYCGLHQTCLR 525  
Db 422 CCRSGKS-----NVGTSGDHDD-----SANKTLRSKMGK-----WCRHCFPC-- 459  
QY 526 LLSYGDPNIIISLQGFALQMGNNENVOQLLOEGISLGNSEADQLLEAAKAGDVETVKKL 585  
Db 460 -----CRGS-----GKSNV-----CASGDHDDSAKMTLR-----NKMKGW 489  
QY 586 CIVQSVNCR-----DIEGRQSTPLHFAAGYNRVSVVEYLLQ-- 621  
Db 490 CCHCFPCRCGSKSKYGAWGDDYDSDAFMEPRYHVRGEDLDKLHRAAWGKVPKRDILVML 549  
QY 622 HGADVHAKDKGLVPLHNACSYGHYEVELLVKHGAVVNVADLWKETPLHEAAAKGKYEI 681  
Db 550 RDTVDNKKQKRTALHLASANGSEVVKLLDRRCOLNVDNKKRTALIKAVQCQDEDEC 609  
QY 682 CKLLQHGADPTKKNRDGNTPLDLVKDGTDIQDILLRGDAAILLDAAKKGLARVKKLSSP 741  
Db 610 ALMLEHGTDPNIPDEYN----- 628  
QY 742 DNVNCRDTCGRHSTPLHLAGAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801  
Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGLVHEQKQV 675  
QY 802 AALLIKYNACVNATDKWAFPTPLHEAAQKGRTOCALLLAHGADPTLKNQBGOTPLDLVSA 861  
Db 676 VKFLIKKANLNALDRYGRATALILAVCCSASIVSLLLEQNDIVSSODLSGQAREYAVS 735  
QY 862 DVVSALLTAAMPSPALPCYKPOVLNGVRSPGATADALSQSPSSLSAASLDNLGSG 921  
Db 736 SHHVI-----CQLLSDYKEKQMLKISSE-----NSNPEQDLKLTSEESQRFKGS 781  
QY 922 FSELSSVSSSGTEGASSLEKKEVPGVDFPSITOFVRNIGLHLMDFERQIITDLVLEM 981  
Db 782 -----ENSQPEKMSQEPENKND-----GDRE-----VE- 804  
QY 982 GHKELKEIGINAYGHRHKLKIGVERLISGOQGLNPVLTNTSGTGTLIDLSPDDKEFQS 1041  
Db 805 --EEMKKHESNNVGLLENLTNGV-TAGNDGNGLIPOKRSRT-----PENQOFPD 850  
QY 1042 VEEMOSTVREHRRDGHAGGIFNRYNLIKQKVCNK-----KLWERYTHRRKEVSEE 1093  
Db 851 NSEEEYHRICE-----LVSDYKEKQPKYSSSENSNPEQDLKLTSEESQRLGSEN 901  
QY 1094 NNNH-----ANERMLFPHGSPFVNAILHKGFE 1120  
Db 902 GOPELENFMAIEEMKKHGS-----THVGFPPE 927

RESULT 10  
US-09-895-814-378  
; Sequence 378, Application US/09895814  
; Publication No. US20020193296A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.



```

: APPLICANT: Hural, John
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Vinals de Bassols, Carlota
: APPLICANT: Foy, Teresa
: APPLICANT: Fanger, Gary R.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
: FILE REFERENCE: 210121.427C26
: CURRENT APPLICATION NUMBER: US/09/895,814
: CURRENT FILING DATE: 2001-06-29
: NUMBER OF SEQ ID NOS: 990
: SOFTWARE: FastSEQ for Windows Version 3.0
: SEQ ID NO 378
: LENGTH: 1719
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-895-814-378

```

Query Match	6.2%;	Score 398.5;	DB 9;	Length 1719;
Best Local Similarity	21.3%;	Pred. No. 1.4e-19;		
Matches 237;	Conservative 137;	Mismatches 411;	Indels 327;	Gaps 34;

Qy	80	CAGGAAACASAAAEVPEAREL-----FEACRNGDVERV-----	KRLVTP	120
Db	72	CRSGKSNVAGSGDHDDSAKMTLRNKGWKWCCHFCPCRGSKGSKYKAGWDYDSDAFMEP	131	
Qy	121	EKVSNRDTAGRKSTPLHFAAGFC---RKDVVEYLLQNGANVOARDGGILPLHNACSFCH	177	
Db	132	R----YVRGEDLDKLRHAAMGKVPVKDLT---VMDRTDVNKKQKKTALHLASANG	185	
Qy	178	AEVYNLLRHGADPNARDNNYTPLHEAAIKGIDYCVILLOHGAEPTIRNTDGRATDLD	237	
Db	186	SEVVKLLDRRCQLNLDNKKRTALIKAVOCQDECALMLEHGTDPNIPDEVGNTTLHY	245	
Qy	238	ADPSAKAVLTGEYKDBELLESARSGNEEKXMMALLTPLNVNCHASDCKSTPLHLAAGYNR	297	
Db	246	A-----IYNEDKLMA-----	255	
Qy	298	VKIVQLLLOHGADVHAKDGLVPLHNACSYGHVEYTELLVHKHGCACVNAMDLWQFTPLHE	357	
Db	256	-----KALLYGADIESKNKHGLTPLLGVHEQOQVVKELIKKANLNALDRYGRATIL	311	
Qy	358	AASKNRVEVCSILLSYGADPTLNCINKNSAIDLAPTP--OLKERTAYEPKGHSLQAAARE	415	
Db	312	AVCCGSASIVSLLEONIDVSSODLSGQTAREYAVSSHVHCQLSDYKEXKQMLKISSE	371	
Qy	416	----ADVTRKKHLSLEMNFRKHPTH---ETALHCAAS---PYPKRQICELLRLKGA	465	
Db	372	NSNPENVSRTR-----NKPRTHMVVEVDSMPAASVVKPGLRSGMKGWCRCRPP	421	
Qy	466	NINEKTEFTPLHVAASEKAHNDVEVVYKHEAKVNALDNLGOTSUHRAAAYCGHJQTCRL	525	
Db	422	CCRESGKS-----NVGTSGDHDD-----SAMKTLRSKMKG-----WCRHCFPC--	459	
Qy	526	LLSYGCDPNIIISLQGTALQMGNEVYQOLLQEBGISLGNSEADRLLEAAKAGDVETVKL	585	
Db	460	-----CRGS-----GKSNV-----GASGDHDDSAKMTLR-----NMGKW	489	
Qy	586	CTVQSYNCR-----DIEGRQSTPLHFAAGYNKVSWEVYLQ--	621	
Db	490	CCHCFCCRGSGSKYKAGWDYDSDAFMEPRYHVRGEDLDKLRHAAMGKVPKDLIVML	549	
Qy	622	HGADVIAHKDGGVPLHNACSYGHVEAYELLVKHGAVVNWADLWKPPLHEAAAAKGYEI	681	
Db	550	RDTDVNKKDKQKRTALHLASANGSEVWKLLDRLRRCQLNVDNKKRTALIKAVOCQDEBC	609	
Qy	682	CKLLLOHGADPTKKNRDNTPDLVDKDGDTDIDQLLRGDAALLDAAKKCLARVKKLSPP	741	
Db	610	ALMLEHGTDPNIPDEYGN-----	628	
Qy	742	DNYNCRDTCGRHSTPLHLAAGYNLNLEVAYELLQHGADVNAQDGGILPLHNAAASYGHVDV	801	



Db 460 -----CRGS-----GKSNV-----GASGDHDSAMKTLR-----NKMKGW 489  
Qy 586 CTQSVNCR-----DIEGROSTPLHFAAGYNRVSVVEYLQ- 621  
Db 490 CCHCFPCRGSGKSKVGAWGYDDSAFMEPRYHVRGEDLDKLHRAAWGKVPKDLIVML 549  
Qy 622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHCAGVNVADLWKFTPLHHEAAAKGYEI 681  
Db 550 RDTDVNKKDKQKRTALHLASANGSEVVKLLDRRCOLNVLDNKKRTALIKAVQOQDEB 609  
Qy 682 CKLLQHGADPTKKNRGTPLDLVKDGDTDIQLLRGDAALDAAKGCLARVKKLSSP 741  
Db 610 ALMLLEHGTDPNIDPEYGN----- 628  
Qy 742 DNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLOHGADVNAODKGLLPLHNAASYGHVDV 801  
Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGLVHEQKQOV 675  
Qy 802 AALLIKYNACVNATDKWFTPLHFAAOKGRTQLCALLAHGADPTLKNQEGOTPLDLIVSA 861  
Db 676 VKFLIKKANLALDRYGRNTALILAVCCGSASIVSLLLEQNIIDVSSQDLSGTAREYAVS 735  
Qy 862 DVSALLTAAMPSPSALPCYKQVNLGVRSPGATADALSSGSPSSLSAASLDNLSGS 921  
Db 736 SHHHVI-----COLLSDYKEKQMKISSE-----NSNPEQDLKLTSEESQRFKGS 781  
Qy 922 FSELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLVEM 981  
Db 851 NESEYHRRICE-----LVSDYKEKQMPKYSSENSNPEQDLKLTSEESQRLGSEN 901  
Qy 1094 NHHN-----ANERMLFHGSPFVNALIIHKGFDE 1120  
Db 902 GQPELENFMAIEEMKKHGS-----THVGFPPE 927  
RESULT 13  
US-09-822-827-378  
; Sequence 378, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.534C1  
; CURRENT APPLICATION NUMBER: US/09/822.827  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 982  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 378  
; LENGTH: 1719  
; TYPE: PRF  
; ORGANISM: Homo sapien  
US-09-822-827-378  
Query Match 6.2%; Score 398.5; DB 10; Length 1719;  
Best Local Similarity 21.3%; Pred. No. 1.4e-19;  
Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;  
Qy 80 CAGGAACASAAAEFAVEPAAREL-----FEACRNGDVERV-----KRLVTP 120  
Db 72 CRGSGKSNVGASGDHDSAMKTLRNMKGKWCCHCFPCCRGSGKSKVGAWGYDDSAFMEP 131  
Qy 121 EKVNSRTAGRKSTPLHFAAGFG---RKDVVEYLLQNGANYQARDGGLPLHNACSGFH 177  
Db 132 R-----YHVRGEDLDKLHRAAWGKVPKDLI--VMLRDTDVNKKDKQKRTALHLASANGN 185

Qy 178 AEVNNLLRHGADPNARDNNWNTPLHEAAIKGKIDVCIVLLOHGAEPTRINTDGR TALDL 237  
Db 186 SEVVKLLDRRCOLNVLDNKKRTALIKAVQOQDEB CALMLLEHGTDPNTPDEYGN TTHY 245  
Qy 238 ADPSAKAVLTGEYKKDELLESARSNGNEEMKMMALUTPLNVNCHASDGRKSTPLHLAAGYNR 297  
Db 246 A-----IYNEDKLMA----- 255  
Qy 298 VKIVOLLQHGADVHAKDKGDLVPLHNACSYGHYEVTTELVKHCAGVNAODLWQETPLHE 357  
Db 256 ----KALLYGADIESKNKHGLTPLLGLVHEQKQOVVFKLKKKANLALDRYGR TALIL 311  
Qy 358 AASKNRVVCVSLLSYGADPTLLNCHNKSADLAPTP--QLKERLAYBFKGHSLQLAARE 415  
Db 312 AVCCGSASIVSLLLEQNIIDVSSQDLSGTAREYAVSSHVVICOLLSDYKEKQMKLISSE 371  
Qy 416 ----ADVTRIKKHLSLEWVNPKHPOTH---ETALHCAAS---PYPKQKQICE LLLRKA 465  
Db 372 NSNPENVSRT-----NKPRTHMVVEVDSMPAASVVKPGLRSKMGKWCRCRFP 421  
Qy 466 NINEKTKEFLTPLHVAESEKAHNDVVVVVHKEAKYNALDNLGOTSILHRAAYCGHLQTCRL 525  
Db 422 CCRSGKS-----NVGTSGDHD-----SAMKTLRSKMGK-----WCRHCFPC-- 459  
Qy 526 LLSYGCDDPNIIISLOQFTALQMGNNVQOLLOEGISLGNSEADROLLEAAKAGDVETVKKL 585  
Db 460 -----CRGS-----GKSNV-----GASGDHDSAMKTLR-----NKMKGW 489  
Qy 586 CTQSVNCR-----DIEGROSTPLHFAAGYNRVSVVEYLQ- 621  
Db 490 CCHCFPCRGSGKSKVGAWGYDDSAFMEPRYHVRGEDLDKLHRAAWGKVPKDLIVML 549  
Qy 622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHCAGVNVADLWKFTPLHHEAAAKGYEI 681  
Db 550 RDTDVNKKDKQKRTALHLASANGSEVVKLLDRRCOLNVLDNKKRTALIKAVQOQDEB 609  
Qy 682 CKLLQHGADPTKKNRGTPLDLVKDGDTDIQLLRGDAALDAAKGCLARVKKLSSP 741  
Db 610 ALMLLEHGTDPNIDPEYGN----- 628  
Qy 742 DNVNCRDTQGRHSTPLHLAGYNNLEVAEYLLOHGADVNAODKGLLPLHNAASYGHVDV 801  
Db 629 -----TTLHYAIYNEDKLMAKALLYGADIESKNKHGLTPLLGLVHEQKQOV 675  
Qy 802 AALLIKYNACVNATDKWFTPLHFAAOKGRTQLCALLAHGADPTLKNQEGOTPLDLIVSA 861  
Db 676 VKFLIKKANLALDRYGRNTALILAVCCGSASIVSLLLEQNIIDVSSQDLSGTAREYAVS 735  
Qy 862 DVSALLTAAMPSPSALPCYKQVNLGVRSPGATADALSSGSPSSLSAASLDNLSGS 921  
Db 736 SHHHVI-----COLLSDYKEKQMKISSE-----NSNPEQDLKLTSEESQRFKGS 781  
Qy 922 FSELSSVSSSGTEGASSLEKKEVPGVDFSTQFVRNLGLEHLMDFEREQITLDVLVEM 981  
Db 782 -----ENSQPEKMSQPEINKD-----GDRE-----VE- 804  
Qy 982 GHKELKEIGINAYGHRHKLIGVERLISGOOGLNPYLTLNTSGSTLIDLSDDKKEFQS 1041  
Db 805 --EEMKKHESNNVGLLENLTNGV--TAGNCDNGLIPQRKSR-----PENQOFPD 850  
Qy 1042 VEEEMQSTVREHRDGGHAGGINFRYNNILKIQVCNK-----KLWERYTHRRKEVSEE 1093  
Db 851 NESEYHRRICE-----LVSDYKEKQMPKYSSENSNPEQDLKLTSEESQRLGSEN 901  
Qy 1094 NHHN-----ANERMLFHGSPFVNALIIHKGFDE 1120  
Db 902 GQPELENFMAIEEMKKHGS-----THVGFPPE 927  
RESULT 14  
US-09-835-788A-12  
; Sequence 12, Application us/09835788A

```
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
; FILE REFERENCE: PT018P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-835-788A-12

Query Match          6.1%; Score 392; DB 10; Length 740;
Best Local Similarity 25.9%; Pred. No. 1.1e-19;
Matches 156; Conservative 71; Mismatches 207; Indels 168; Gaps 15;

QY 39 RALSASFGLGALLAGPG-LLRLALLLAVAAARIMSGRCAGGGAAC----- 87
Db 249 QGLLVQEPDG--LWVATPAQTLTDTLDDLIAAVSTRVPTGSSNSQTECLTPESCQTT 306
QY 88 ASAAAEAVEPA-----ARELFEACRNGDVRVKRLVTPPE-KVNSRQDTAGRK 132
Db 307 SNVASQSMPPVPSVDIDAHTESNHDLTALTACAGGHEELVSLIARDAKIEHRDKG-- 364
QY 133 STPLFAAGFRKDVVYVLLQNGANVQARDGGL-IFLHNACSFHGAEVNLLLRHGDAP 191
Db 365 FTPLILAATAGHGVVVEILLDKGDIQAQSERTKDPLSLACSGGROEVVDLLARGANK 424
QY 192 NARDNNYTPLHBAIKGKIDVICIVLQHGAEPTIRNTDGTALDLDAPSAAVLTGEYK 251
Db 425 EHRNVSQDYTPLSLAASGGYVNIKIILLNAGAE----- 456
QY 252 KDELLESARSGNEEKMMALITPLNVNCHASDGRKSTPLHLAAGYNRVKIYVOLLLOHCAOV 311
Db 457 -----INSRTGSKLGISPLMLAAMNGHVPAYKLLLDMGSDI 492
QY 312 HAK-DKGDVPLHNACSYGHEVTELLVKKHGACVNAMDLMQFTPLHEAASKNRVEVCSSL 370
Db 493 NAQIETNRNTALTACFOGRAEVVSLLLDRKANVEHRAKTGLTPLMEASGGYAEVGRVL 552
QY 371 LSYGADPTLLNCHKNKAIDLAFTPOLKERLAYEFKGHSLLOAAREADVTRIKKHLSEMV 430
Db 553 LDKGAD-----VNAPPV----- 564
QY 431 NFKHPOTHETALHCAASPYPRKQICELLRLKRGANINEKTKBELTPLHVASEKAHNDVV 490
Db 565 -----PSRDTALTIAADKGHYK---FCELLIHGCAHIDVRKNGTNPWLASNGGHFDVV 617
QY 491 EVVVKHEAKVNALDNIQOTSLHRAAYCGHLQTCRLLS 528
Db 618 QLLVQAGADVDAADNRKRTPLMSAFRKGHVKVVOYLKVEVNFQPSDIECMRYIATITDKE 677
QY 529 -----YCCDPNIISLOGFTALMGNNVQOLLOE-GISLGNSEADQLLEAAKAGDVETV 582
Db 678 LLKKHOCVETIVKAKDQAAE-ANKNASILLKELDLEKSRRESRQALAAKREKRKR 736
QY 583 KK 584
Db 737 KK 738

RESULT 15
US-09-947-199-2
```

```
; Sequence 2, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-947-199-2

Query Match          5.1%; Score 332.5; DB 10; Length 835;
Best Local Similarity 26.1%; Pred. No. 2.2e-15;
Matches 111; Conservative 60; Mismatches 124; Indels 131; Gaps 14;

QY 422 KKHLSEMVNFKHP-----QTHETALHCAASPYPRKQICELLRLKRGANINEKTKBELTLP 477
Db 81 KSHIRTLMLKGLRPSRLTRNGFTALHLAV---YKDNAELITSLHSGADIQQVGYGGLTA 137
QY 478 LHVASEKAHNDVVVEVYVVKHEAKVNALDNIQOTSLHRAAYCGHLQTCRLLSYGCDPNLI 537
Db 138 LHATITAGHLEAADVLLQHGANNVQIDAVFTPLHIAAYYGHEQVTRLLKFGADVNV-- 195
QY 538 LOGFTALMGNNVQOLLOEGISLGNSEADQLLEAAKAGDVETVKKLCTVQSVNCRDIE 597
Db 196 -----SGEVGD----- 201
QY 598 GRQSTPLHFAAGYNRVSVVEYVLLQHG--ADVHAKDKGLVPLHNACSYGHEVAELLVKH 655
Db 202 -----RPLHLASAKGFLNIAKLLMEEGSKADVNQADNEDHVPFLHPCSRFGHHDIIVKYLQS 257
QY 656 GAVV--NVADLWKTPLHEAAAKGYEICKLLLO-HGADP-TKKNRDGNT----- 701
Db 258 DLEVQPHVNNIYGDTPHLACYNGKFEVAKETIIQISGESLTLENIFSETAFHSACTYK 317
QY 702 PLDLVKDGDTDIODLLRGDAALLDAAKKGCLARVKLLSSPDNVNCRDTQGRHS-TPLHLA 760
Db 318 SIDLVK-----FLLD-----QNVINIHQGRDGTGHLHA 347
QY 761 AGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 820
Db 348 CYHGHIRLVQFLDNGADMN-----LVACDPKSKSGEKDEQTCMLM-----WAY 390
QY 821 TPLHEA 826
Db 391 EKGHDA 396

Search completed: February 12, 2003, 07:53:05
Job time : 31.4359 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:07:14 ; Search time 24.3761 Seconds  
(without alignments)  
4890.316 Million cell updates/sec

Title: US-09-843-159B-4  
Perfect score: 6464  
Sequence: 1 RCSARRGAAGGGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	895	13.8	1765	2 T42714	ankyrin 3, splice
2	895	13.8	1961	2 T42716	ankyrin 3, splice
3	892.5	13.8	4377	2 A55575	ankyrin 3, long sp
4	891	13.8	1943	2 T42713	ankyrin 3, splice
5	889	13.8	1940	2 T42715	ankyrin 3, splice
6	880.5	13.6	1856	2 B35049	ankyrin 1, erythro
7	880.5	13.6	1880	2 A35049	ankyrin 1, erythro
8	878.5	13.6	1881	1 SJHUK	ankyrin 1, erythro
9	877	13.6	3924	2 S37431	ankyrin 2, neuroha
10	876	13.6	1848	2 S37771	ankyrin, erythrocy
11	874	13.5	1862	2 T49502	ankyrin - mouse
12	866.5	13.4	1549	2 T13940	ankyrin - fruit fl
13	858	13.3	2039	2 T15347	ankyrin-related un
14	657	10.2	791	2 T42691	hypothetical prote
15	563.5	8.7	934	1 H71374	probable ankyrin -
16	552	8.5	1411	2 S30355	alpha-latroinsecto
17	507.5	7.9	1058	2 D82654	ankyrin-like prote
18	494.5	7.7	1062	2 T30255	inversin - mouse
19	489.5	7.6	1062	2 T14151	Inv protein - mous
20	485	7.5	1401	2 S11527	alpha-latrotoxin p
21	463	7.2	1188	2 T19552	hypothetical prote
22	452.5	7.0	426	2 AE2149	hypothetical prote
23	426	6.6	991	2 T25412	hypothetical prote
24	416.5	6.4	1184	2 T00253	gene Ankhzn protei
25	399	6.2	2584	2 T24158	hypothetical prote
26	399	6.2	2606	2 T24157	hypothetical prote
27	389.5	6.0	397	2 T46445	hypothetical prote
28	370	5.7	1031	2 T43458	hypothetical prote
29	350	5.4	557	2 T46507	hypothetical prote

30 348.5 5.4 606 2 AC2508 hypothetical prote  
31 345 5.3 1423 1 T37275 death-associated p  
32 343 5.3 1107 2 T21280 hypothetical prote  
33 329.5 5.1 1435 2 T32930 hypothetical prote  
34 321 5.0 1398 2 T21884 hypothetical prote  
35 315.5 4.9 656 2 A34793 sex-determining pr  
36 307 4.7 1004 2 A51142 myosin-light-chain  
37 306 4.7 1016 2 T19006 ankyrin related pr  
38 306 4.7 1322 2 A59288 myosin heavy chain  
39 305 4.7 679 2 B45771 2-5A-dependent RNA  
40 299.5 4.6 633 2 T27499 hypothetical prote  
41 299 4.6 247 2 D84448 probable ankyrin l  
42 297 4.6 1001 2 S30385 G9a protein - huma  
43 297 4.6 1295 2 T21720 hypothetical prote  
44 294.5 4.6 900 2 A42024 transcription fact  
45 285.5 4.4 658 2 S68418 protein phosphatas

RESULT 1  
T42714  
ankyrin 3, splice form 2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C:Accession: T42714  
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eichler, E.M.; Higgins, A.; Yialamas, M.; Turtzo  
J. Cell Biol. 130, 313-330, 1995  
A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin ge  
the repeat domain.  
A:Reference number: 222237; MUID:95340633; PMID:7615634  
A:Accession: T42714  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1765 <PEP>  
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1  
A:Experimental source: strain C57BL/6J; kidney  
C:Genetics:  
A:Gene: Ank3  
A:Map position: 10  
A:Introns: 1587/1  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing

Query Match 13.8%; Score 895; DB 2; Length 1765;  
Best Local Similarity 30.3%; Pred. No. 1.6e-46;  
Matches 294; Conservative 134; Mismatches 409; Indels 134; Gaps 25;

Qy 68 AVAAARIMSGRRRCAGGGACASAAAAEVEPAARELFACRNGDVERKRLVTPKVNSRD 127  
Db 10 AKPAHRKRGKGSNDANASYLRAARAGHLEKA----LDYIKNG-----VDVNICN 54  
Qy 128 TAGRKSTPLHFAAGGRKDVVEYLQNGANVQARDGGILPLHNACSFGEAEVWLLRH 187  
Db 55 QNGLNA--LHLASKEGHVEVSELLQREANVDAATKKGNTALHLIASLAGAEVVKVLVN 112  
Qy 188 GADPNARDNNWYTPHAAIKGKIDVCIVILQHGAEPTIRNTDGTALDLA-----DFSA 242  
Db 113 GANVNAQSQNGFTPLYMAAQENHLEVRFLDNGASQSLATEDGFTPLAVALQQGHQDV 172  
Qy 243 KAVLTGEYK---KDBLLESARSNGNEEKMMALLTPLNVNCHASDGRKSTPLHAAAGYNRVK 299  
Db 173 SLLLENDTKGVRLPALHIAARKDDTKAAALLQNDTNADVESKSGFTPLHIAAHYGNIN 232  
Qy 300 IVQLLLQHGADVHAKKGDVLPPLHNACSYGHEVTELLVKHGACVNMADLMWQFTPLHEAA 359  
Db 233 VATLLNRAAAVDFTARNIDITPLHVASKRGANVMVKLLLDRCAGKIDAKTRDGLTPLHCGA 292  
Qy 360 SKNRVEVCSLLSYGADFTLLNCHNKSALDLAPTQPKERLAYEFKSHSLQAAAEADV 419  
Db 293 RSGHEQVVEMLLDRSA-PILSKTKN-----GLSPLHMAQGD-- 328

ALIGNMENTS





```
Db 1084 AELLNGMEELDSPEEL--GTRKRICRIITKDFPO-----YFAVVSRIKQES 1127
QY 1143 NOYVYGIGGG 1152
Db 1128 NQ--IGPEGG 1135

RESULT 5
T42715
ankyrin 3, splice form 3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42715
R:Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Vialamas, M.; Turtzo
J. Cell Biol. 130, 313-330, 1995
A:Title: Anky3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A:Reference number: 222237; MUID:95340633; PMID:7615634
A:Accession: T42715
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1940 <PEP>
A:Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1
A:Experimental source: strain C57BL/6J; kidney
C:Genetics:
A:Gene: Ank3
A:Map position: 10
A:Introns: 834/1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing

Query Match 13.8%; Score 889; DB 2; Length 1940;
Best Local Similarity 27.9%; Pred. No. 4.4e-46;
Matches 337; Conservative 168; Mismatches 496; Indels 206; Gaps 38;

QY 68 AVAAARIMSGRRGAGGACACASAAAEVPAARELPACRNGDVERKLVTEKVNRSR 127
Db 10 AKPAHRRKRGKKSDANASYLRAARAGHLEKA---LDYIKNG-----VDVNICN 54
QY 128 TAGRKSTPLHFAAGFGKDVVEYLQNGANVQARDGGLPLHNACSFHGAHVNNLLRH 187
Db 55 QNGENA--LHLASKEGHVEVVSSELLQREANVDAATKGNALHIALAGAEVVKVLVN 112
QY 188 GADPNARDNNWYTPLEHAAIKGIDVICIVLQHGAEPTIRTDGRTALDLA-----DPSA 242
Db 113 GANVNAQSQNGFTPLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQGHQDVV 172
QY 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
Db 173 SLLENDTKGKVLPAALHIAARKDDTKAAALLLONDNDVADVESKSGFTPLHIAAHYGNIN 232
QY 300 IVOLLQHGADVHAKDKGDLPLHNACSYGHVEVTELLVKGACVNAWDLWQFTPLHEAA 359
Db 233 VATLLNRAAAVDFTARNDITPLHVASKRGNANVKKLLLDGAKIDAKTRDGLTPLHCGA 292
QY 360 SKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTPOLKRLAYEFKGHSLLOQAAREADVT 419
Db 293 RSGHEQVEMVLEDRSA--PILSKTKN-----GLSPLHMATQGD-- 328
QY 420 RIKKHLISLEMVNFKH-----PQTHETALHCAASAPYPRKQICELLRLKGNINIEKTK 472
Db 329 ----HUNCVQLLQHNVPVDDVTNDYLTALHVAACHGCHK---YAKVLLDKKASPNKAL 381
QY 473 EFTPLHVAASEKANDVVEVVVHKAQVNDNLGQTSLHRAAYCGLHQTCLRLLSYGCD 532
Db 382 NGFTPLHIACKKRIIRVMELLKHGASIQAVTESGLTPIHVAAFMGHVNIQSOLMHGAS 441
QY 533 PNIIISQGTALQMG-----NENVOQLQBGISIGNSEADRG--LLEAAKAGDVETVKKL 585
Db 442 PNTNVRGETALHMAARSGOAEVVRVLDQGAQVEAKAKDDQTPHLISARLGKADIVOQL 501
QY 586 CTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLQHGADVHAKDKGLVPLHNAACSYG 644

Db 55 QNGENA--LHLASKEGHVEVVSSELLQREANVDAATKGNALHIALAGAEVVKVLVN 112
QY 188 GADPNARDNNWYTPLEHAAIKGIDVICIVLQHGAEPTIRTDGRTALDLA-----DPSA 242
Db 113 GANVNAQSQNGFTPLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQGHQDVV 172
QY 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
Db 173 SLLENDTKGKVLPAALHIAARKDDTKAAALLLONDNDVADVESKSGFTPLHIAAHYGNIN 232
QY 300 IVOLLQHGADVHAKDKGDLPLHNACSYGHVEVTELLVKGACVNAWDLWQFTPLHEAA 359
Db 233 VATLLNRAAAVDFTARNDITPLHVASKRGNANVKKLLLDGAKIDAKTRDGLTPLHCGA 292
QY 360 SKNRVEVCSLLSYGADPTLLNCHNKSAIDLAPTPOLKRLAYEFKGHSLLOQAAREADVT 419
Db 293 RSGHEQVEMVLEDRSA--PILSKTKN-----GLSPLHMATQGD-- 328
QY 420 RIKKHLISLEMVNFKH-----PQTHETALHCAASAPYPRKQICELLRLKGNINIEKTK 472
Db 329 ----HUNCVQLLQHNVPVDDVTNDYLTALHVAACHGCHK---YAKVLLDKKASPNKAL 381
QY 473 EFTPLHVAASEKANDVVEVVVHKAQVNDNLGQTSLHRAAYCGLHQTCLRLLSYGCD 532
Db 382 NGFTPLHIACKKRIIRVMELLKHGASIQAVTESGLTPIHVAAFMGHVNIQSOLMHGAS 441
QY 533 PNIIISQGTALQMG-----NENVOQLQBGISIGNSEADRG--LLEAAKAGDVETVKKL 585
Db 442 PNTNVRGETALHMAARSGOAEVVRVLDQGAQVEAKAKDDQTPHLISARLGKADIVOQL 501
QY 586 CTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLQHGADVHAKDKGLVPLHNAACSYG 644
Db 502 LOQASPNAAATTSQ--YTPHLAAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAYG 559
QY 645 HYEVAELLVKGAVVNVADLWKTTPLEHAAKCKEYELCKLLLOHGADPTKKNRDNTPLD 704
Db 560 KLEVASLLQKSPDAAAGSGTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLH 619
QY 705 L-VKDGDTQIDQLL-----RGDAALDAAKKGCLARVKKLSSPD-VNVCRTQ 750
Db 620 IAAKNQMDIATSELLEGADANAVTQGTASVHLAQEGHVDVMSVLLSRNANVNSNKS 679
QY 751 GRHSTPLHLAAGYNNLEVAEYVLLQHGADVNAQDKGLPLHNAASYGHVDVAALLIKYNA 810
Db 680 G--LTPHLAAQEDRVNAEVLNQAHDVDAQTKMGYTPPLHVGCHYGNIKIVNFFLQHSA 737
QY 811 CVNATDKWATPLHEAAQKRTOLCALLLAHGADPTLKNOEGQTPDLNVS-----ADD 863
Db 738 KVNAKTNGYTAHQAAQOQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYSVVD 797
QY 864 VSALLTAAMPSPALPSYK---POVLN-----GVRSPGAT-----AD 897
Db 798 LKVVTETIWTITTEKHKNVPETNNEVLDMSDDEVRKASAPKLSGDEYISDGEGED 857
QY 898 ALSS-----GP-----SSPSSLSAASSLDNLGSEFSELSVVVS-----SSQTEGA 937
Db 858 AITGDTDKYLGPDQLKELGDDSLPAEGYVGFSLGARSASLRSFSSDRSYTLNRSYARDS 917
QY 938 SSLEKKEVPGVDSITQFVNIGLHMDIFEREQITLDVL-----VENGHKELKEIGI 991
Db 918 MMEIFELVPSKEGHLT--FTREFSDSLRH--YSWAADTLDNVLNVLSSPVHSGF--LVSFMV 973
QY 992 NAYG-----HRHKL-----KQVERLISQOQGLNPILTINTSGSTILIDLS 1033
Db 974 DARGSGMRGRHGMRIIIPRKTAPTITCRLVRKHKLANPPVWEGEGSLASRLVEMG 1033
QY 1034 PDKKQFOS-VEEQMSTVREHROGGHAGGTINRYNLIKIOKVNKKLWE--RYTHRRKEV 1090
Db 1034 PAGAQLGLPVIVETP-----HFGSMRGKEREILVLSRNETGETWKEHQFDSKNEDL 1083
QY 1091 SE-----ENHNANERMLPHGSPFVNATIHKGFDERHAYIGGMFAGAGIYFAENS---SKS 1142
```



Db 502 LOOGASPNATTS--YTPHLHAAAREGHDVAFLLDHGCASISITTKGFTPLHVAAYG 559  
QY 645 HYEVALLKHGAVNVADLWFTPLHEAAAKGYEICKLLLOHGADPTKKRNGDNTPLD 704  
Db 560 KLEVASLLLOKSPDAACKSGLTPLHVAHYDNOKVALLLDQAGSPHAAKNGYTPPLH 619  
QY 705 L-VKGDGTDIODLL-----RGDAALLDAKAGKGLARVKKLSSPD-NVNCRTDQ 750  
Db 620 IAAKKNMIDATSLLEYGADANAVTROGTASVHLAAQEGHDMVSLLSRNANVLSNKS 679  
QY 751 GRISTPLHAAAGYNNLEVEAYLLOHGADVNAQDKGLIPLHNAASYGHVDAALLIKNA 810  
Db 680 G--LTPHLHAAQDRVNAEVLVNOGHVDAQTKMGTPLHVGCHYKTIKIVNFFLLQSHA 737  
QY 811 CVNATDKWFTPLHEAAQGRQTQALCALHAGADPTLKNQEGQTPLDLVS-----ADD 863  
Db 738 KVNAKTKNGYTALHQAQOQHTHIINVLONNASPNELTVNGTALAIARRLGIYSVDVT 797  
QY 864 VSALLTAAPPPSALPSCYK---PQVLNGVRSPG-----ATADALS 900  
Db 798 LKVVTEETITTTITEKHKNMNPETMNEVLDMSDDGDKCTWFKIPKVOEVLVKSEDAIT 857  
QY 901 S-----GP-----SPSSLSAASSLDNLGSELSFSSVVS-----SSGTECASSL 940  
Db 858 GDTYKYLGPQDLKELGDDSLPAGYVGFSLGARSASLSRFSFSDRSYTLNRSYARDSMMI 917  
QY 941 EKKEVPQVDFISITQFVRNLGLEHLMDFEREQITLDVL-----VEMGHKELKEIGINAY 994  
Db 918 EELLVPSKEOHLT-FTRFEDSDSLR-YSWAADTLDNVNLVSSPVHSGF--LVSEMVDA 973  
QY 995 G-----HRHKL-----KGVRLISGGQGLNPYTLTNVSGSGTILDLSPDD 1036  
Db 974 GSGMRGSRHGMRIIPPRKCTAPTTRITCLVKRHLANPPMPVEGEGLASRLVEMGPAG 1033  
QY 1037 KEFQS-VEEEMQSTVREHRODGHAGGIFNRYNLTKQKCNKLEW--RYTHRRKEVSP- 1092  
Db 1034 AQPLGPVIVEP-----HFGSMRGKRELIVLRSENGETWKEHQFDSKNEDLAEL 1083  
QY 1093 -----ENHNHANRMLFHGSPFVNAIHKGFDERHAYIGMGAGIYFAENS---SKSNQY 1145  
Db 1084 LNMDEELDSPEEL---GTRICRIITKDFPQ-----YFAVVSRIKQESNQ- 1126  
QY 1146 VYIGGG 1152  
Db 1127 -IGPEGG 1132

RESULT 6  
B35049  
ankyrin 1, erythrocyte splice form 3 - human  
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R  
N:Contains: ankyrin 2.2, erythrocyte  
C:Species: Homo sapiens (man)  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 10-Jul-1998  
C:Accession: B35049  
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, M.C.; K  
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990  
A:Title: cDNA sequence for human erythrocyte ankyrin.  
A:Reference number: A35049; MUID:90175370; PMID:1689849  
A:Accession: B35049  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1856 <LAM>  
C:Genetics:  
A:Gene: GDB:ANK1; ANK  
A:Cross-references: GDB:118737; OMIM:182900  
A:Map position: 8p11.2-8p11.2  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing  
F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>  
F:2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>  
F:44-76/Domain: ankyrin repeat homology <AN01>  
F:77-109/Domain: ankyrin repeat homology <AN02>

F:110-142/Domain: ankyrin repeat homology <AN03>  
F:143-171/Domain: ankyrin repeat homology <AN04>  
F:172-204/Domain: ankyrin repeat homology <AN05>  
F:205-237/Domain: ankyrin repeat homology <AN06>  
F:238-270/Domain: ankyrin repeat homology <AN07>  
F:271-303/Domain: ankyrin repeat homology <AN08>  
F:304-336/Domain: ankyrin repeat homology <AN09>  
F:337-369/Domain: ankyrin repeat homology <AN10>  
F:370-402/Domain: ankyrin repeat homology <AN11>  
F:403-435/Domain: ankyrin repeat homology <AN12>  
F:436-468/Domain: ankyrin repeat homology <AN13>  
F:469-501/Domain: ankyrin repeat homology <AN14>  
F:502-534/Domain: ankyrin repeat homology <AN15>  
F:535-567/Domain: ankyrin repeat homology <AN16>  
F:568-600/Domain: ankyrin repeat homology <AN17>  
F:601-633/Domain: ankyrin repeat homology <AN18>  
F:634-666/Domain: ankyrin repeat homology <AN19>  
F:667-699/Domain: ankyrin repeat homology <AN20>  
F:700-732/Domain: ankyrin repeat homology <AN21>  
F:733-765/Domain: ankyrin repeat homology <AN22>  
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match 13.68; Score 880.5; DB 2; Length 1856;  
Best Local Similarity 26.08; Pred. No. 1.4e-45;  
Matches 294; Conservative 156; Mismatches 374; Indels 307; Gaps 29;

QY 105 ACRNGDVERVRLVTPKVNRSRDYAGRKSTPLHFAAGFGRKDVVEYLQNGANGYQARDG 164  
Db 52 ASKEGHVKKVVELLHKE-IILETTTKGNTALHIAALAGQDEVVRELNVYNAVNAQSOK 110  
QY 165 GLIPLHNACSPGHARVWMLLRHGHADPNARDNNWYTPHHEAAIKGKIDVCTVLQHGAE 224  
Db 111 GFTPLMAAENHLEVVVKFLLENGANQVATEDGFTPLAVALQOQGHENVAHLINYGTKG 170  
QY 225 TI-----RNTDGRALDLADPSAKAVLTGEYKDELLESARSGNEERKMALLPLN 275  
Db 171 KVRPLALHIAARNDTRTA-----AVLLNDPNPDL--SKTG-----FTPLH 211  
QY 276 VNCHASD-----GRKS-----TPHLAAGYNNRYKIYOLLQHGADVHAKDG 317  
Db 212 IAAHYENLNAQVLLNRGSSVNFTPQGITPLHIAARRGNVIMVRLLDGQAETKTD 271  
QY 318 DLVPLHNACSVGHYEVTTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSELLSYGADP 377  
Db 272 ELTPLHCAARNGHVRISBELLDHGAPIQAKNGIUSPTHMAAQGDHDCVRLLLQYDAE- 330  
QY 378 TLLNCHNKSADLAPTPOKLERLAYEFKGHSLQAAREADVTRIKKHLSELVNFKHPQT 437  
Db 331 -----ID-----DIT-----LD 337  
QY 438 HETALHCAAAPYKPKQICELLRLKGANINEKTEFTPLHVAESEKAHNDVVEVVKHE 497  
Db 338 HLTPHLVAA---HCGHHRVAKVLLDKGAKPNRSLALNGFTPLHIAACKNHRVMEILLKTG 394  
QY 498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLOGFTALONG-----NENVO 552  
Db 395 ASIDAVTESGLTPLHVASFPMGHLPVTKNLLORGASPNVSNKVPETPLHMAARAGHTEYAK 454  
QY 553 QLLQE-----GISLGNSEADROLLE-----AAKAG 577  
Db 455 YLLQNAKVNAKAKDDQTPHCAARIHGHTNMVKKLLENANPNLATTAGTPLHIAAREG 514  
QY 578 DYETVKLCTVQ-SYNCRDIEGRQSTPLHFAAGYNNRVSVVEYLQHGADVHAKDGGLVP 636  
Db 515 HVETVLALLEKEASQACMTKKG--FTPLHVAAYKVKRVAELLERDAHPNAAAGNGLTP 572  
QY 637 LHNACSVGHYEVAELLVKGAVNVADLWK-FTPLHEAASKGYEICKLLQHGADPTKK 695  
Db 573 LHVAVHHNLDIVKLLPRGSGSPH-SPAWNGYTPHIAAKQNVQEVARSLLQYGSANAE 631  
QY 696 NRDGNTPLDL-VKGDGTDIQLDGLRGAALLDAAKGLARVKKLSPPDNVNCRDTOGRHS 754  
Db 632 SVOGVTPLHAAQEGHAEWVALL-----LSKQANGNLGNKSG--L 669

[illegible]





Db 603 AAHYDNOKVALLLEKASPHATAKNGYTPLHIAAKKNQMOIASTLLNNGAETNIVTKQG 662  
Qy 700 NTPDLV-KDGDTDIQLDRLGDAALLDAAKKGCGLARVKKLSSPDNVNCRDTQGRHSTPLH 758  
Db 663 VTPLHLSQEGHTDMVTLKDKGNIHMTKSLG-----TSLH 700  
Qy 759 LAAGYNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818  
Db 701 LAQEDKVNVAIITKHGADDAHTKLGYPPLIVACHYGNVMVFNLLKOGANNAKTKN 760  
Qy 819 AFTPLHEAAOKGRTQLCALLAHGADPTLKNQEGOTPLDL-----VSADDVSALLTAAM 872  
Db 761 GYTLHQAQOQHHTIINVLQHGAKGNATTANGNTALAIKRLGYISVDTLKVVTEEV 820  
Qy 873 PPSALPCYKQVNLGKVRSPGATADALSSGSPSSLSAASLNLGSGFSEL----- 925  
Db 821 TTTTITTEK-----HKLNVPTMTVEV-----DYSDEGGDTMTGCGEYLRPEDLK 868  
Qy 926 ----SSVVSSTGEGASLEKKEVPGVDFSTQFV--RNGLGLEHMLDIFEREQITLVLV 979  
Db 869 ELGDDSLPSSQFLDGMNLYRYSLEGGRSDSLRSFSDRSHTLSHAS--YLDRSAMMDSDV 926  
Qy 980 EMGHKELKEIGINAYGHRHKLKIGVERL 1007  
Db 927 VIPSHQVSTLAKEARNYRLSWGTENL 954  
RESULT 10  
S3771  
ankyrin, erythrocyte - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 27-May-1994 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999  
C:Accession: S3771  
R:Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.  
J. Biol. Chem. 268, 9533-9540, 1993  
A:Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an  
A:Reference number: S3771; MUID:93252825; PMID:8486643  
A:Status: preliminary  
A:Accession: S3771  
A:Molecule type: mRNA  
A:Residues: 1-1848 <BIR>  
A:Cross-references: EMBL:X59063; NID:g311816; PIDN:CA448801.1; PID:g311817  
C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing  
F:48-80/Domain: ankyrin repeat homology <AN01>  
F:81-113/Domain: ankyrin repeat homology <AN02>  
F:114-146/Domain: ankyrin repeat homology <AN03>  
F:147-175/Domain: ankyrin repeat homology <AN04>  
F:176-208/Domain: ankyrin repeat homology <AN05>  
F:209-241/Domain: ankyrin repeat homology <AN06>  
F:242-274/Domain: ankyrin repeat homology <AN07>  
F:275-307/Domain: ankyrin repeat homology <AN08>  
F:308-340/Domain: ankyrin repeat homology <AN09>  
F:341-373/Domain: ankyrin repeat homology <AN10>  
F:374-406/Domain: ankyrin repeat homology <AN11>  
F:407-439/Domain: ankyrin repeat homology <AN12>  
F:440-472/Domain: ankyrin repeat homology <AN13>  
F:473-505/Domain: ankyrin repeat homology <AN14>  
F:506-538/Domain: ankyrin repeat homology <AN15>  
F:539-571/Domain: ankyrin repeat homology <AN16>  
F:572-604/Domain: ankyrin repeat homology <AN17>  
F:605-637/Domain: ankyrin repeat homology <AN18>  
F:638-670/Domain: ankyrin repeat homology <AN19>  
F:671-703/Domain: ankyrin repeat homology <AN20>  
F:704-736/Domain: ankyrin repeat homology <AN21>  
F:737-769/Domain: ankyrin repeat homology <AN22>  
F:770-802/Domain: ankyrin repeat homology <AN23>  
Query Match 13.6% Score 876; Db 2; Length 1848;  
Best Local Similarity 29.7% Pred. No. 2.6e-45;  
Matches 250; Conservative 132; Mismatches 281; Indels 178; Gaps 21;  
Qy 105 ACRNGDVERVKRLVTPKVNSRDTAGRKSTPLHFAAGGRKDVVEYLLQNGANVQARDG 164

Db 56 ASKEGHVKNVVELLHKE-IILETTTKGNTALHIAALAQDEWRELVNNGANVAQSQK 114  
Qy 165 GLIPLHNACSGHAENVNLLRHGADPNARDWNWYTPLHEAIAKIDVICIVLLOHGAEP 224  
Db 115 GFTPLYMAAQENHLEVVKFLLENGANQVATEDGFTPLAVALQOQHENVVAHLINYGTKG 174  
Qy 225 TI-----RNTDRTALDLADPSAKAVLTGEYKDELESARSNEKEMMALITPLN 275  
Db 175 KVRPLALHIAARNDDTFTA-----AVLLQNDPNPDVL--SKTG-----FTPLH 215  
Qy 276 VNCHASDGRKS-----TPHLAAGYNRVKIVOLLQHOGADVHAKDKG 317  
Db 216 IAAHYENLVNAQLLNRGASVNFTPQNGITPLHIASSRGVIMVRULLDRGAQIETRTKD 275  
Qy 318 DLVPLHNACSGHYEVTELLVKHGACVAMDLWQFTPLHEAASKNRVVCSSLLSYGADP 377  
Db 276 ELTPLHCAARNGHVRISEILDHGAPIQAKYKNGUSPLHMAAQGDHLDVRLULQYNAE- 334  
Qy 378 TLLNCHNKSAIDLAPTQPKERLAYEFKGHSLLOAREADVTRIKKHLSLEMVNFKHPT 437  
Db 335 -----ID-----DIT-----LD 341  
Qy 438 HETALHCAAASPYPKRKOICEALLLRKANINEKTKFEFTPLHVASEKAHNDVVEVVKHE 497  
Db 342 HLTPLHVAA--HCGHHRVAKVLLDKGAPNSRALNGFTPLHIACKKNHIRVWELLKTKG 398  
Qy 498 AKVNALDNLGOTSLHRAAYCGHLOTCRLLLSYGCDPNIIISLOGFTALQM----GNENVQO 553  
Db 399 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 458  
Qy 554 LIQEGISLGNSEAD-----ROLLE-----AAKAG 577  
Db 459 YLLONKAKANAKAKDDQTPHCAARIGHTGVKVLLENLNGASPNLATTAGTPTLHTAAREG 518  
Qy 578 DVETVKILCTVO-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGLVP 636  
Db 519 HVDYALALLEKEASQACMTKKG--FTPLHVAAYKGVRLAEELLEHDHAHPNAAGKNGLTP 576  
Qy 637 LHNACSGHYEVAELLVKHGAVNVVADLWK-FTPLHEAAYKGYEICKLLLOHGADPKK 695  
Db 577 LHVAVHHNLDIVKULLPRGGSPH-SPAWNGYTPLHIAAKQNIQIIVARSLLQYGGSANAE 635  
Qy 696 NRDGNTPLDL-VKGDGTDIQLDRLGDAALLDAAKKGCGLARVKKLSSPDNVNCRDTQGRHS 754  
Db 636 SVQGYTPPLHAAQEGHTEWALL-----LSKQANGNLGNKSG--L 673  
Qy 755 TPLHLAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNA 814  
Db 674 TPLHLVSQEGHVPVADVLIKHGVTVDAITRMGYTPLHVASHYGNIKLVKFLQHQADVNA 733  
Qy 815 TDKWAFTPLHEAAOKGRTQLCALLAHGADPTLKNQEGOTPLDL-----VSADDVSALL 868  
Db 734 KTKLGYSPHQAQOQHHTDITVTLKNGASPNVSNNGTTPLAIAKRLGYISVTVLKV 793  
Qy 869 T 869  
Db 794 T 794  
RESULT 11  
I49502  
ankyrin - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 13-Aug-1999  
C:Accession: I49502  
R:White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.  
Mamm. Genome 3, 281-285, 1992  
A:Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory  
A:Reference number: I49502; MUID:92343717; PMID:1386265  
A:Accession: I49502  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1-1862 <RES>  
A:Cross-References: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940  
C:Genetics:  
A:Gene: Ank-1  
C:Superfamily: ankryrin; ankryrin repeat homology  
C:Keywords: alternative splicing  
F:40-72/Domain: ankryrin repeat homology <AN01>  
F:73-105/Domain: ankryrin repeat homology <AN02>  
F:106-138/Domain: ankryrin repeat homology <AN03>  
F:139-167/Domain: ankryrin repeat homology <AN04>  
F:168-200/Domain: ankryrin repeat homology <AN05>  
F:201-233/Domain: ankryrin repeat homology <AN06>  
F:234-266/Domain: ankryrin repeat homology <AN07>  
F:267-299/Domain: ankryrin repeat homology <AN08>  
F:300-332/Domain: ankryrin repeat homology <AN09>  
F:333-365/Domain: ankryrin repeat homology <AN10>  
F:366-398/Domain: ankryrin repeat homology <AN11>  
F:399-431/Domain: ankryrin repeat homology <AN12>  
F:432-464/Domain: ankryrin repeat homology <AN13>  
F:465-497/Domain: ankryrin repeat homology <AN14>  
F:498-530/Domain: ankryrin repeat homology <AN15>  
F:531-563/Domain: ankryrin repeat homology <AN16>  
F:564-596/Domain: ankryrin repeat homology <AN17>  
F:597-629/Domain: ankryrin repeat homology <AN18>  
F:630-662/Domain: ankryrin repeat homology <AN19>  
F:663-695/Domain: ankryrin repeat homology <AN20>  
F:696-728/Domain: ankryrin repeat homology <AN21>  
F:729-761/Domain: ankryrin repeat homology <AN22>  
F:762-794/Domain: ankryrin repeat homology <AN23>

Query Match 13.5%; Score 874; DB 2; Length 1862;  
Best Local Similarity 29.7%; Pred. No. 3.5e-45;  
Matches 250; Conservative 132; Mismatches 281; Indels 178; Gaps 21;

QY 105 ACRNGDVERVKRLVTEPKVNSRDTAGRKSTPLHFAAGFGKDWVEYLLQNGANVQARDG 164  
Db ASKEGHVKMVELLHKE--IILETTTKKGTALHIAALAGODEVRELNYGANVNAQSK 106

QY 165 GLIPLNACSFHAEVYVLLRHGADPNARDNNYTPLEHAALKGKIDVICVILQHCAPP 224  
Db GTTPLYMAAQENHLEVVKFLLLENGANQNVATEDGFTPLAVALQOQHENVVAHLINVTGK 166

QY 225 TI-----RNTDGTALDADPSAKAVLTGEYKDELLESARGNEEKMMALLTPN 275  
Db KVLRLPALHIAARNDDRTA-----AVLLQNDPNPDVL--SKTG-----FTPLH 207

QY 276 VNCHASDGRKS-----TPLHLAGYNRVKIVQLLLOHGADVHAKDG 317  
Db IAHYENLVNQALLNRGASVFTPOGITPLHIAARRGNVIMVRLLDGRGAQIETRKD 267

QY 318 DLVPLNACSYGHEVTELLVKHGACVNAIDLWQFTPLHFAAASKNRVEVCSLLLSYGADP 377  
Db EUTPLHCAARNHVRISSEILLDHGAPLOAKTNGLSPIHMAAQGDHLCVRLLLQYNAE- 326

QY 378 TLLNCHNKSAIDLAPTPOLKERLAYEFKGHSLLQAAAREADVTRIKHLSLEVMVNFKHPT 437  
Db ID-----DIT-----LD 333

QY 438 HETALHCAASYPKPKQICELLRLKRGANINEKTEFLPLHVASEKAHNDVVEVVKKE 497  
Db HLTPLHVA--HCGHRRVAKLLDQKAPNSRALNGFTPLHIAKKNHIRMVLELLKRG 390

QY 498 AKVNALNDLGOTSLHRAAYCGHLOTCLRLLSYCGDNPITLSOGFTALOM---GNENVOQ 553  
Db ASIDAVTESGLTPLHVASFWGHLPIVKNLLQRGASPNVSVKVEPLHMAARAGHTEVAK 450

QY 554 LLOEGTSLNSEAD-----ROLLE-----AAKAG 577  
Db YLLQNKAKANAKAKDDQTPHCAARITGTMVKVLLLENGASPNLATTAGHTPLHTAAREG 510

QY 578 DVETVKKLTQV--SYNCRDIEGRQSTPLHFAAGYNEVSVVEYLLQHGADVHAKDGGLVP 636  
Db HVDTALLLEKASQCMTKG--FTPLHVAAYKGVRLAELELLEHDAHPNAAKGNGLTP 568

QY 637 LHNACSYGHEVYAEELLVKHGAVVNVADLMK-FTPLHFAAASKYETCKILLQHGADPTKK 695  
Db LHVAVHHNNLDIVKLLPRGSGPH-SPANNGYTPPLHIAAKQNIQVARSLLQYGGSANAE 627

QY 696 NRDGNTPDL-VKGDGTDIDQLLRGDAALLDAAKCKLARVKKLSPPDNVNCRDTOGRHS 754  
Db SVQGVTPPLHAAQEGHTEMVALL-----LSKQANGNLGNKSG--L 665

QY 755 TPLHLAGYNNLEVAEYLLQHGADVNAODKGLIPLHNAASYGHVDVAALLIKYNACVNA 814  
Db TPLHVSQEGHVLVADVLIKHGVTVDATTRMGYTPPLHVASHYGNIKLVKFLQHQADVNA 725

QY 815 TDKWAFTPLHFAAQAQKGTQCALLLAHGADPTLKNOEGOTPLDL-----VSADYSALL 868  
Db KTKLGYSPHLQAQOQGHDTDIVLLKNGASPNVSSNGTTPPLAIAKRLGLVISVTDVLKVV 785

QY 869 T 869  
Db T 786

RESULT 12  
TI3940  
ankryrin - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
C:Accession: TI3940  
R:Dubreuil, R.R.; Yu, J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994  
A:Title: Ankryrin and beta-spectrin accumulate independently of alpha-spectrin in Dros  
A:Reference number: Z17820; MUID:95024098; PMID:7937942  
A:Accession: TI3940  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1549 <DUB>  
A:Cross-References: EMBL:L35601; NID:g557083; PID:g557084; PIDN:AAAC37208.1  
C:Genetics:  
A:Cross-References: FlyBase:FBgn0011747

Query Match 13.4%; Score 866.5; DB 2; Length 1549;  
Best Local Similarity 27.4%; Pred. No. 7.5e-45;  
Matches 311; Conservative 173; Mismatches 492; Indels 159; Gaps 34;

QY 98 AARELFEACRNGDVERVKRLV---TPEKVNSRDTAGRKSTPLHFAAGFGKDWVEYLLQ 154  
Db ATISFLRAARSGDIKKVMDFLDCGEISDINSNANGINA--LHLAAKDGVDICCELLRR 95

QY 155 GANVQARDGGIPLNACSFHAEVYVLLRHGADPNARDNNYTPLEHAALKGKIDVC 214  
Db GIKIDNATKKGNTALHTASLAGQHDVINQLLYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLQHGAEPTTNTDGTALDLA-----DPSAKAVLTGEYK---KDELLESARSNEEK 266  
Db RTLLANGANPSLSTEDGFTPLAVAMQOQGHDKIVAVLELNDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALTPLVNCHASDGRKSTPLHLAGYNRVKIVQLLLOHGADVHAKDGDLVPLHNA 326  
Db AKLLQLQHPNADIVSKSGFTPLHIAAHYGNVDIATLLNKNKADVNVVAKHNITPLHVAC 275

QY 327 SYGHEVTELLVKHGACVNAIDLWQFTPLHFAAASKNRVEVCSLLLSYGADPTLLNCHN-K 385  
Db KWKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLQONA-PILTCKNGL 334

QY 386 SAIDLAPTOLKERLAYEFKGHSLLQAAAREADVTRIKHLSLE-----MVNF 432  
Db SALHMAAQE-----HDEAAHILLDNKAPVDEVTDYLTALHVAACHGKVKVAKLLDY 388

QY 433 K-HPOTHE---TALHCAASYPKPKQICELLRLKRGANINEKTEFLPLHVASEKAHN 487  
Db KANPNARALNGFTPLHIACKK---NRKVVELLIKGANIGATTESGLTPLHVASFWGCI 445

QY 488 DVVEVVVKHEAKVNALDNLGOTSLHRAAYCGHLOTCLRLLSYCGDNPITLSOGFTAL--- 544

```
Db 446 NIVYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDALVREGOTPLHVA 504
QY 545 -QMGEN-VQOLLEGLISGNSEADR--QLLEAAKAGDVETVKKLCIVQSVNCRDIEGRQ 600
Db 505 SRLGNIINIIMLLQHGAEINNAQSNKYSAHIAAKEGOENIVQVLLNGAEN-NAVTKKG 563
QY 601 STPLHEAAGNVRVSVYVLLQHGADVHAKDKGGLVPLHNACSYGHEVAEVLVKGAVVN 660
Db 564 FTPLHLACKYGVQNVQVLLQNGASIDFGKNDVTPPLHVATHYNPISIVELLKANGSPN 623
QY 661 VADLWKFTPLHEAAAKYKEICKLLQHGADPTKKNRDNTPDL-VKDGDTDIDQLLRG 719
Db 624 LCARNQCATHIACKKNYLEAMQLLQHGADVNIISKSGFSPHLAAQGGNVDMVQLLE 683
QY 720 DAALLDAKKCLARVKLLSPDNVNCRDGTGRSTPLHLAAGYNUNLEVAEYLLQHGADV 779
Db 684 YGVLSAAAKNGL-----TFLHVAQAQEGHVLVSQILLEHGANI 720
QY 780 NAQDKGGLIPLHNAASYGHVDAALLIKYNACVNAATDKWFTPLHEAAQKRTOLCALLL 839
Db 721 SERTNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAQOQHIMIINLLL 780
QY 840 AHGADPTLKNQOGTPLDVSADVSALLTAAMPSPALPSCYKPOVLNG-----V 889
Db 781 RHKANPNALTKDGN TALHASN-----LGYVTVMESLKI VITSTSVINSIGATEERLKV 834
QY 890 RSPCATADALSSGSPSSLSAAS-----SLDNLGSGFSFELSSVSSSGTE-----GAS 938
Db 835 MTPELMGETLLSDSDDESCDLDLHNYKYMATDLDLANYQDQKNFTTNTDHLTDVS 894
QY 939 SLEKKEVPGVDFITQFVRNLGLEHMD--TFEREQITLDLVEMGHKELKEIGINAYGH 996
Db 895 VLNKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF---LYDARGSGMRGY 948
QY 997 RHKLKIGVE-----RLISQOGLNPLYLTNTSGSGLTILDLSPDDKEFQS-- 1041
Db 949 RHNGVRIITVPPKACAEPTITRCYVKVQPVVNPPLMEGEALVSKILEMSPVDMGFLSPI 1008
QY 1042 -VEEMOSTVREHRDGGHAGIFRNYNLLKQVCNKKLWERYTHRRKEVSEENHNHANE 1100
Db 1009 TLEVPHYGTLRK-----NEREII-ILRSNDGESWREH-NLYKDIIGEDINQFEE 1055
QY 1101 RMLFHGSPFVNAIIHKGDERHAYIGGMFGAGIVFAENSSKNQVYVYIG--GGT 1153
Db 1056 --FHSDRIVR-IVTQNPVPHFAVV-----SVRQEVHVIGPDGGT 1092

RESULT 13
T15347
ankyrin-related unc-44 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Aug-2002
C:Accession: T15347; T15346; T15344; T15345; A57282; B57282; C57282
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid B0350.
A:Reference number: Z18332
A:Accession: T15347
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2039 <GAT>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208873; PIDN:AAA93443.1
A:Accession: T15346
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1000,'SKLOHRT',1002-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVAES',1956-1957,'EQ
32','S',2034-2035,'GSPTRRSVEPEEHRHSHOHEGST' <GA2>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1
A:Accession: T15344
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1718,'KW',1903-1905,'NRLADESSPS',1916-1917,'QRSTIVAESTSEQVPE',1934-1935,'B
```

```
<GA3>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208874; PIDN:AAA93444.1
A:Accession: T15345
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVAES',1956-1957,'EQ
PTRRSVEPEEHRHSHOHEGST' <GA4>
A:Cross-references: EMBL:U50071; NID:g1208871; PID:g1208876; PIDN:AAA93446.1
R:Otsuka, A.J.; Franco, R.; Yang, B.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakul
J. Cell Biol. 129, 1081-1092, 1995
A:Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in
A:Reference number: A57282; MUID:95263663; PMID:7744957
A:Accession: A57282
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-852,'GGG',856-1000,'SKLOHRT',1002-1319,'IG',1322-1595,'DA',1598-1718,'K
','SHRED',2007-2008,'TI',2011,2017,'TT',2020-2022,'SHIS' <OTS>
A:Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608
A:Accession: B57282
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 831-852,'GGG',856-1319,'IG',1322-1595,'DA',1598-1718,'KWEELNRL',1727,1799
V',1945-1947,'VT',1950,'SH',1975,'SESP',1980-1981,'SPTRRSVEPEEHRHS',1984-1985,'EDHEGS
A:Cross-references: GB:U21731
A:Accession: C57282
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 194,'E',196,'I',198,'DC',201,'G',409,'AV',412,'Q',414,824,'S',826,'ER',82
4,'TIV',1828,'ESTS',836,'QV',1836,'E',1934-1935,'EOS',1939,'ESES',1944,'REDDGFTIVTT',
A:Cross-references: GB:U21732; NID:g790603; PIDN:AAA85853.1; PID:g790604
C:Genetics:
A:Gene: CESP:unc-44
A:Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1
C:Superfamily: ankyrin; ankyrin repeat homology
F:164-192/Domain: ankyrin repeat homology <AN04>
F:358-390/Domain: ankyrin repeat homology <ANI>
F:391-423/Domain: ankyrin repeat homology <AN11>

Query Match 13.3%; Score 858; DB 2; Length 2039;
Best Local Similarity 26.3%; Pred. No. 3.9e-44;
Matches 319; Conservative 163; Mismatches 443; Indels 286; Gaps 39;

QY 91 AAEEVEPAARE---LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db 22 APAAPEPEAGESASFLRAARAGDLEKLELLRAGCTDINTSNAGLNS--LHLASKEGHS 79
QY 146 DVVEYLLONGANVQARDGGGLIPLHNACSFHAEVNVNLLRHGADPNARDNNWYTPLHEA 205
Db 80 EYVRELIRKQAVDAATRKNGTALHIASLAGOSLIVTILVENGANVNVQSVNGFTPLYMA 139
QY 206 ATKGKIDVICVILQHGAEPTIRNTDGTALDLADPSAKAVLTGEYKDELLESARSNPE 265
Db 140 AQENHEEVVYKLLKHGANAQALSTEDGFTPLAVA-----LQOQHDR 179
QY 266 KMWALLPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLQHGADVHAKDKGLVPLHN 324
Db 180 VVAVLLE-----NDSKGVRLPALHIAAKDKDTTAATLLQNEHNPDPVTSKSGFTPLHI 233
QY 325 ACSYGHYEVTTELLVKHGACVNMALWQFTPLHEAAKSNRVEVCSSLLSYGADPTLLNCHN 384
Db 234 AAHYGHENVQQLLEKGANVNYQARHNISPLHVATKWGTNMANLLLSRGA---IIDSRT 290
QY 385 KSAIDLATTPOLKERLAYEFKGHSLLQAREADVTRIKKHSLEVMVNFKHPQTHETALHC 444
Db 291 KDLL-----TPLHC 299
QY 445 AAASPYPRKQTCICELLRRKANINEKTEFLTPLHVASEKAHNDVVVVVKKHAKVN--A 502
Db 300 AARSGH---DQVOLLVVQGAIPISAKTKNGLAPLHMAAGDHVDAARTLLYHRAPVDVY 356
QY 503 LDNLGQTSLHRAAYCGHLOTCELLLSYGCOPNIISLQGFALOMGNEN----- 550
Db 357 VDYLL--TPLHVAACHGVHVRVAKLILLDRSDPNRSALINGFTPLHIACKKNRKIVKVELLKY 414
```









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 21:15:54 ; Search time 15.3675 Seconds  
(without alignments)  
3346.711 Million cell updates/sec

Title: US-09-843-159B-4

Perfect score: 6464

Sequence: 1 RCSARRGAGGQAGRGARY.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6108	94.5	1166	1	TNK2_HUMAN
2	5103.5	79.0	1327	1	TNK1_HUMAN
3	892.5	13.8	4377	1	ANK3_HUMAN
4	878.5	13.6	1880	1	ANK1_HUMAN
5	877	13.6	3924	1	ANK2_HUMAN
6	874	13.5	1862	1	ANK1_MOUSE
7	760.5	11.8	1059	1	Y379_HUMAN
8	571	8.8	768	1	YB23_HUMAN
9	485	7.5	1401	1	LATA_LATMA
10	431	6.7	603	1	V162_FOWPV
11	399.5	6.2	747	1	V222_FOWPV
12	390.5	6.0	832	1	ANK3_HUMAN
13	382.5	5.9	596	1	V024_FOWPV
14	380.5	5.9	668	1	V244_FOWPV
15	357	5.5	587	1	ANK2_HUMAN
16	354.5	5.5	776	1	ANK5_HUMAN
17	354.5	5.5	1431	1	DAPK_HUMAN
18	351	5.4	542	1	V155_FOWPV
19	351	5.4	775	1	ANK5_MOUSE
20	327.5	5.1	525	1	V228_FOWPV
21	322	5.0	525	1	ANK3_MOUSE
22	321	5.0	436	1	V245_FOWPV
23	318	4.9	592	1	V246_FOWPV
24	315.5	4.9	656	1	PEML_CAEEL
25	312.5	4.8	568	1	PPGB_MOUSE
26	310	4.8	528	1	PPGA_HUMAN
27	309	4.8	524	1	PPGA_MOUSE
28	308	4.8	518	1	ANK3_HUMAN
29	307.5	4.8	568	1	PPGB_BOVIN
30	305	4.7	735	1	RN5A_MOUSE
31	304	4.7	567	1	PPGB_HUMAN
32	303	4.7	429	1	AS10_HUMAN
33	300.5	4.6	692	1	ANK6_HUMAN

34	291.5	4.5	741	1	RN5A_HUMAN
35	291.5	4.5	950	1	ORP1_HUMAN
36	285	4.4	461	1	V218_FOWPV
37	285	4.4	1083	1	Y1L2_YEAST
38	283.5	4.4	226	1	PSDA_HUMAN
39	281.5	4.4	434	1	V219_FOWPV
40	281.5	4.4	482	1	V232_FOWPV
41	281.5	4.4	898	1	KBF2_HUMAN
42	280.5	4.3	231	1	PSDA_RAT
43	278	4.3	984	1	KBF1_CHICK
44	277	4.3	1964	1	NTC4_MOUSE
45	275	4.3	347	1	GABC_MOUSE

ALIGNMENTS

RESULT 1  
ID TNK2\_HUMAN STANDARD: PRT; 1166 AA.  
AC Q9H2K2; Q9HASA; Q9H8F2;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-interacting ankyrin-related App-ribose polymerase 2) (Tankyrase-like protein) (Tankyrase-related protein).  
DE TNKS2 OR TNKL OR TANK2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=21072550; PubMed=11205898;  
RA Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W., Glass B., Meese E.U.;  
RA "Novel tankyrase-related gene detected with meningioma-specific sera";  
RT Clin. Cancer Res. 7:113-119(2001).  
RL [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Breast Carcinoma;  
RX MEDLINE=21190090; PubMed=11294570;  
RA Kulmov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K., Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;  
RT "Cloning and characterization of TNKL, a member of tankyrase gene family";  
RL Genes Immun. 2:52-55(2001).  
RN [3]  
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.  
RC TISSUE=Liver;  
RX MEDLINE=21264473; PubMed=11278563;  
RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M., Eyre H.J., Sutherland G.R., Daly R.J.;  
RT "Identification of a novel human tankyrase through its interaction with the adaptor protein Grb14";  
RL J. Biol. Chem. 276:17172-17180(2001).  
RN [4]  
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCALIZATION.  
RC TISSUE=Placenta;  
RX MEDLINE=21443728; PubMed=11454873;  
RA Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D., Morin G.B., Yaswen P., Campisi J.;  
RT "TANK2, a new TRF1-associated poly(Adp-ribose) polymerase, causes rapid induction of cell death upon overexpression";  
RL J. Biol. Chem. 276:35891-35899(2001).  
RN [5]  
RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND LNPEP/OTASE.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=21661461; PubMed=11802774;



Db 481 LOEGTSLGNSEADROLLAANKAGDVETVKLCTQVSVNCRDIEGRQSTPLHFAAGYNRYVS 540  
 QY 615 VVEYLLOHGADVHAKDKGLVPLHNACSYGHVEAELLVKGAVVNVADLWKFTPLHEAA 674  
 Db 541 VVEYLLOHGADVHAKDKGLVPLHNACSYGHVEAELLVKGAVVNVADLWKFTPLHEAA 600  
 QY 675 AKGYEICKLLOHGADPTKKNRQDNTPLDLVKDGDITDQDLLRGDAALLDAARKGCLAR 734  
 Db 601 AKGYEICKLLOHGADPTKKNRQDNTPLDLVKDGDITDQDLLRGDAALLDAARKGCLAR 660  
 QY 735 VKLSSPDVNCRDQGRHSTPLHLAAGYNVLEVAEYLLQHCADVNAQDKGLIPLHNA 794  
 Db 661 VKLSSPDVNCRDQGRHSTPLHLAAGYNVLEVAEYLLQHCADVNAQDKGLIPLHNA 720  
 QY 795 SYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAQKGRTOCALLLAHGADPTLKNQEGQT 854  
 Db 721 SYGHVDVAALLIKYNACVNAATDKWAFTPLHEAAQKGRTOCALLLAHGADPTLKNQEGQT 780  
 QY 855 PLDLVSADVDVSAALLTAAMPSPALPSCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 914  
 Db 781 PLDLVSADVDVSAALLTAAMPSPALPSCYKPOVLNGVRSFGATADALSSGSPSSLSAASS 840  
 QY 915 LDNLGSGFSELSSVSVSSGTEGASLEKKEVPGVDFTSIQFVRLNGLBHLMDIFEREQIT 974  
 Db 841 LDNLGSGFSELSSVSVSSGTEGASLEKKEVPGVDFTSIQFVRLNGLBHLMDIFEREQIT 900  
 QY 975 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSGTILIDLSP 1034  
 Db 901 LDVLVEMGHKELKEIGINAYGHRHKLKIGVERLISGQOGLNPYLTLNTSGSGTILIDLSP 960  
 QY 1035 DDKFQSVSEEMQSTVREHRDGGHAGGTFNRYNLIKQKVCNKKLWERYTHRRKEVSEEN 1094  
 Db 961 DDKFQSVSEEMQSTVREHRDGGHAGGTFNRYNLIKQKVCNKKLWERYTHRRKEVSEEN 1020  
 QY 1095 HNANERMLFHGSPFVNAILHKGDERHAYIGGMFGAGIYFAENSSKNQVYIGGGTG 1154  
 Db 1021 HNANERMLFHGSPFVNAILHKGDERHAYIGGMFGAGIYFAENSSKNQVYIGGGTG 1080  
 QY 1155 CPVHKDRSCYICHRQLLCFRVTLGKSLQFSAMKMAHSPPGHSHSVTGRPSVNGLALAEYV 1214  
 Db 1081 CPVHKDRSCYICHRQLLCFRVTLGKSLQFSAMKMAHSPPGHSHSVTGRPSVNGLALAEYV 1140  
 QY 1215 IYRGEQAYPEYLITYQIMRPGMVDG 1240  
 Db 1141 IYRGEQAYPEYLITYQIMRPGMVDG 1166

RESULT 2  
 TNK1\_HUMAN STANDARD; PRT; 1327 AA.  
 ID TNK1\_HUMAN  
 AC O95271; O95272;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tankyrase 1 (EC 2.4.2.30) (TNK1) (tankyrase 1) (TNKS-1) (TRF1-  
 interacting ankyrin-related ADP-ribose polymerase).  
 GN TNKS OR TNKSI OR TIN1 OR TINFI OR PARPL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Testis;  
 RX MEDLINE=99040105; PubMed=9822378;  
 RA Smith S., Giriat I., Schmitt A., de Lange T.:  
 RT "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";  
 RL Science 282:1484-1487 (1998).  
 RN [2]  
 RP SUBCELLULAR LOCALIZATION.  
 RX MEDLINE=99454782; PubMed=10523501;  
 RA Smith S., de Lange T.:  
 RT "Cell cycle dependent localization of the telomeric PARP, tankyrase,

to nuclear pore complexes and centrosomes.";  
 J. Cell Sci. 112:3649-3656 (1999).  
 [3]  
 FUNCTION, AND PHOSPHORYLATION.  
 MEDLINE=20556282; PubMed=10988299;  
 Chi N.-W., Lodish H.F.:  
 RT "Tankyrase is a golgi-associated mitogen-activated protein kinase  
 substrate that interacts with IRAP in GLUT4 vesicles.";  
 J. Biol. Chem. 275:38437-38444 (2000).  
 [4]  
 FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.  
 MEDLINE=21602874; PubMed=11739745;  
 Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.:  
 RT "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2  
 at human telomeres.";  
 Mol. Cell. Biol. 22:332-342 (2002).  
 CC -!- FUNCTION: May regulate vesicle trafficking and modulate the  
 subcellular distribution of SLC24A4/GLUT4-vesicles. Has PARP  
 activity and can modify TRF1, and thereby contribute to the  
 regulation of telomere length.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor =  
 nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.  
 CC -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with  
 the cytoplasmic domain of LNPP/otase in SLC24A4/GLUT4-vesicles.  
 CC -!- Binds to the N-terminus of telomeric TRF1 via the ANK repeats.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and  
 with juxtanuclear SLC24A4/GLUT4-vesicles. A minor proportion is  
 also found at nuclear pore complexes and around the pericentriolar  
 matrix of mitotic centrosomes. During interphase, a small fraction  
 of TNKS is found in the nucleus, associated with TRF1.  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be  
 produced by alternative splicing.  
 CC -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.  
 CC -!- PTM: Upon insulin-stimulation, phosphorylated on serine residues  
 by MAPK kinases.  
 CC -!- PTM: ADP-ribosylated (-auto).  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -!- SIMILARITY: CONTAINS 15 ANK REPEATS.  
 CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AF082556; AAC79841.1; -  
 EMBL; AF082557; AAC79842.1; -  
 EMBL; AF082558; AAC79843.1; -  
 EMBL; AF082559; AAC79844.1; -  
 DR HSSP; Q00420; IAWC.  
 DR Genew; HGNC:11941; TNKS.  
 DR MIM; 603303; -  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR001660; SAM.  
 DR Pfam; PF00023; ank; 21.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 15.  
 DR SMART; SM00454; SAM; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 15.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50105; SAM\_DOMAIN; 1.  
 DR Transferrase; Glycosyltransferase; NAD; Golgi stack; Telomere;  
 Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;  
 phosphorylation; Alternative splicing.  
 FT REPEAT 215 247 ANK 1.  
 FT REPEAT 248 280 ANK 2.  
 FT REPEAT 281 313 ANK 3.  
 FT REPEAT 368 400 ANK 4.  
 FT REPEAT 401 433 ANK 5.

```

FT REPEAT 434 466 ANK 6.
FT REPEAT 521 556 ANK 7.
FT REPEAT 557 589 ANK 8.
FT REPEAT 590 622 ANK 9.
FT REPEAT 683 715 ANK 10.
FT REPEAT 716 748 ANK 11.
FT REPEAT 749 781 ANK 12.
FT REPEAT 836 868 ANK 13.
FT REPEAT 869 901 ANK 14.
FT REPEAT 902 934 ANK 15.
FT DOMAIN 1030 1089 SAM.
FT DOMAIN 1176 1327 PARP.
FT DOMAIN 9 14 POLY-HIS.
FT DOMAIN 27 34 POLY-PRO.
FT DOMAIN 128 134 POLY-SER.
FT DOMAIN 137 145 POLY-SER.
FT VARSPIC 641 643 EST -> GHS (IN ISOFORM 2).
FT VARSPIC 644 1327 MISSING (IN ISOFORM 2).
FT MUTAGEN 1184 1184 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED WITH A-1291.
FT MUTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED WITH A-1184.
FT SEQUENCE 1327 AA; 142010 MW; E14DE985C710B957 CRC64;

Query Match 79.0%; Score 5103.5; DB 1; Length 1327;
Best Local Similarity 79.2%; Pred. No. 1.6e-299;
Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

QY 22 AAGTAPDVTAGSOARALSASSPGGLALLAGPGLLRLLRLLALLLAVAAARTMSGRCA 81
DB 112 SAAGVAPNPAGSGSNSSPSSTSSSSSSPSPO-----SSLAESPEAAGVSTAPL 165
QY 82 GGAACASAAAEEAVEPAAREFACRNGDVERVYKRLVTPPEKYNSTRDTAGKSTPLHFAAG 141
DB 166 GPAAAGTGVPVAVSGALRELEACRNGDVSRYKRLVDAANVAKMDACKRKSPLHFAAG 225
QY 142 FGRKDVVELLQNGANVQARDGGLIPLHNACSFHAEVNVLLLRHGADPNARDNNYTP 201
DB 226 FGRKDVVELLQNGANVHARDGGLIPLHNACSFHAEVNVLLLRHGADPNARDNNYTP 285
QY 202 LHEAAIKGIDVICVLLQHGAEPTIRNTDGTALDADPSAKAVLTGEYKKDELLESARS 261
DB 286 LHEAAIKGIDVICVLLQHGADPNIRNTDGTALDADPSAKAVLTGEYKKDELLESARS 345
QY 262 GNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLVP 321
DB 346 GNEEKWALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGDLVP 405
QY 322 LHNACSYGHYEVELLIVKHGACVANDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLN 381
DB 406 LHNACSYGHYEVELLIVKHGACVANDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLN 465
QY 382 CHNKAIDLAPTQPKERLAYEPKSHLSLQAAAREADVTRIKKHLSEVMNFKHPQTHETA 441
DB 466 CHKSAMDAPTPELRLETYEPKSHLSLQAAAREADLAKVKTALALEINFQKQPSHETA 525
QY 442 LHCAASAPPKRQKQICELLRLKGANINEKTKPELTPLHVAASEKAHNDVVEVVKHKAUN 501
DB 526 LHCAVASLHPKRRQKQVTELLRLKGANINEKTKPELTPLHVAASEKAHNDVVEVVKHKAUN 585
QY 502 ALDNLGQTSILHRAAYGHLQTCRLLLSYGDPNITISLOGFTALQMGNEVQQLLQEGISL 561
DB 586 ALDTLGTALHRAALAGHLQTCRLLLSYGDPSIISLQGTAAQMGNEAVQQLLSSTPTI 645
QY 562 GNSEADRLLEAAKAGDVETVKKLTCTVQSVNCRDIEGROSTPLHFAAGYNRVSVVEYLLQ 621
DB 646 RTSVDVYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705
QY 622 HGADVHAKDKGGLVPLHNACSYGHYEVELLIVKHGAVNVVADLWKTPLHAAAKGYEI 681
DB 706 HGADVHAKDKGGLVPLHNACSYGHYEVELLIVKHGAVNVVADLWKTPLHAAAKGYEI 765
QY 682 CKLLQHGADPTKKNRDNTPDLVRDGDPTDIDLLRGDAALLDAACKGCLARVKKLSSP 741

```

```

DB 766 CKLLKHGADPTKKNRDNTPDLVRDGDPTDIDLLRGDAALLDAACKGCLARVKKLCTP 825
QY 742 DNVNCRDPTGHRSTPLHLAAGYNVLEAYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
DB 826 ENINCRDPTGHRSTPLHLAAGYNVLEAYLLQHGADVNAQDKGGLIPLHNAASYGHVDI 885
QY 802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTOQLCALLLAHGAADPTLKNQEGOTPLDLVSA 861
DB 886 AALLIKYNACVNATDKWAFTPLHEAAQKGRTOQLCALLLAHGAADPTMKNQEGOTPLDLATA 945
QY 862 DVSALLTAAMPSPALPSCYKQVNLGVNVRSPGATADALSSGSSPSSLSAASLDNLSSG 921
DB 946 DDIRALLIDAMPPEALPTCFKQAT-----VVSASLISPASTPSCLSAASIDNLTP 998
QY 922 FSELSVSVSSGTEGASSLEKK--EVPGVDEFTQVRNLGLEHLMDFEREQITLIDLVLV 979
DB 998 LAELAVGASNAGDGAAGTERGEVAGLDNMISQFLKSLGLEHLDIPETQITLIDLVA 1058
QY 980 EMGHKELKEIGINAYGHRHKLKINGVERLISGQOGLNPYLTNTSGSGTILIDLSPDKEF 1039
DB 1059 DMGHEELKEIGINAYGHRHKLKINGVERLISGQOGLNPYLTNTSGSGTILIDLSPDKEF 1118
QY 1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNAN 1099
DB 1119 QSVEEEMQSTVREHRDGGHAGGIFNRYNLIKQVCNKKLWERYTHRRKEVSEENHNAN 1178
QY 1100 ERMFLHGSFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKNQYVYGGTGGCPVHK 1159
DB 1179 ERMFLHGSFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKNQYVYGGTGGCPVHK 1238
QY 1160 DRSCYICHRQLLFCRVTLGKSLQFQFAMKMAHSPGHSHSVTGRPSVNGALARYIYRGE 1219
DB 1239 DRSCYICHRQLLFCRVTLGKSLQFQFAMKMAHSPGHSHSVTGRPSVNGALARYIYRGE 1298
QY 1220 QAYPEYLITYQIMRPE 1235
DB 1299 QAYPEYLITYQIMRPE 1314

RESULT 3
ANK3_HUMAN
ID ANK3_HUMAN STANDARD; PRT; 4377 AA.
AC Q12955;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 3 (ANK-3) (Ankyrin G).
GN ANK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem;
RX MEDLINE=95138209; PubMed=7836469;
RA Kordeli E., Lambert S., Bennett V.;
RT "Ankyrin3. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier.";
RL J. Biol. Chem. 270:2352-2359(1995).
CC -!- FUNCTION: Membrane-cytoskeleton linker.
CC -!- ALTERNATIVE PRODUCTS: A number of isoforms are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: Expressed in brain and other tissues.
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

```



RN [3]  
 RP VARIANT HS ILR-462.  
 RX MEDLINE-96225450; PubMed-8640229;  
 RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,  
 RA Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,  
 RA Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;  
 RT "Ankyrin-1 mutations are a major cause of dominant and recessive  
 RT hereditary spherocytosis.";  
 RL Nat. Genet. 13:214-218(1996).  
 CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL  
 CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO  
 CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE  
 CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.  
 CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE  
 CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;  
 CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC  
 CC PLASMA MEMBRANE.  
 CC -!- ALTERNATIVE PRODUCTS: At least 3 isoforms; 1/2.1 (shown here),  
 CC 2/2.2 and 3; are produced by alternative splicing.  
 CC -!- PTM: REGULATED BY PHOSPHORYLATION.  
 CC -!- PTM: ACYLATED BY PALMITIC ACID GROUP(S).  
 CC -!- DISEASE: Defects in ANK1 are the cause of dominant and recessive  
 CC hereditary spherocytosis (HS).  
 CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.  
 CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X16609; CAA34610.1; -.  
 DR EMBL; M28880; AAA51732.1; -.  
 DR PIR; S08275; SJHUK.  
 DR PIR; A35049; A35049.  
 DR HSSP; Q00420; LAWC.  
 DR Genew; HGNC:492; ANK1.  
 DR MIM; 182900; -.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00023; ank; 24.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00791; ZU5; 1.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 22.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00218; ZU5; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 20.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat;  
 KW Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;  
 KW Polymorphism.  
 FT INIT\_MET 0 0  
 FT DOMAIN 1 826 89 KDA DOMAIN (ANION EXCHANGE PROTEIN  
 FT BINDING DOMAIN).  
 FT DOMAIN 827 1381 62 KDA DOMAIN (SPECTRIN BINDING  
 FT DOMAIN).  
 FT DOMAIN 1382 1880 55 KDA REGULATORY DOMAIN (REGULATES  
 FT THE BINDING OF ANKYRIN TO SPECTRIN  
 FT AND THE BAND 3 PROTEIN).  
 FT REPEAT 43 72 ANK 1.  
 FT REPEAT 76 105 ANK 2.  
 FT REPEAT 109 138 ANK 3.  
 FT REPEAT 142 171 ANK 4.  
 FT REPEAT 173 200 ANK 5.  
 FT REPEAT 204 233 ANK 6.  
 FT REPEAT 237 266 ANK 7.

FT REPEAT 270 299 ANK 8.  
 FT REPEAT 303 332 ANK 9.  
 FT REPEAT 336 365 ANK 10.  
 FT REPEAT 369 398 ANK 11.  
 FT REPEAT 402 431 ANK 12.  
 FT REPEAT 435 464 ANK 13.  
 FT REPEAT 468 497 ANK 14.  
 FT REPEAT 501 530 ANK 15.  
 FT REPEAT 534 563 ANK 16.  
 FT REPEAT 567 596 ANK 17.  
 FT REPEAT 600 629 ANK 18.  
 FT REPEAT 633 662 ANK 19.  
 FT REPEAT 666 695 ANK 20.  
 FT REPEAT 699 728 ANK 21.  
 FT REPEAT 732 761 ANK 22.  
 FT REPEAT 765 794 ANK 23.  
 FT DOMAIN 1402 1486 DEATH.  
 FT VARSPPLIC 1512 1873 MISSING (IN ISOFORM 2).  
 FT VARSPPLIC 1874 1874 H -> D (IN ISOFORM 2).  
 FT VARSPPLIC 1849 1880 TWEGPLEDSELEVDIDYPMKHSKDHSTPNP -> ELRGS  
 FT 3).  
 FT 3).  
 FT 3).  
 FT R -> T.  
 FT /FTId=VAR\_000595.  
 FT V -> I (IN HS).  
 FT /FTId=VAR\_000596.  
 FT R -> H (IN BRUEGGEN).  
 FT /FTId=VAR\_000597.  
 FT V -> A.  
 FT /FTId=VAR\_000598.  
 FT D -> E.  
 FT /FTId=VAR\_000599.  
 FT E -> D.  
 FT /FTId=VAR\_000601.  
 FT S -> T.  
 FT /FTId=VAR\_000600.  
 FT D -> N (IN DUESSELDORF).  
 FT /FTId=VAR\_000602.  
 FT R -> D.  
 FT /FTId=VAR\_000603.  
 FT A -> S (IN REF. 2).  
 FT CONFLICT 229 229 V -> I (IN REF. 2).  
 FT CONFLICT 1545 1545  
 SQ SEQUENCE 1880 AA; 206145 MW; 1C5F5E7EFDICD428 CRC64;  
 Query Match 13.6%; Score 878.5; DB 1; Length 1880;  
 Best Local Similarity 25.8%; Pred. No. 8.2e-45;  
 Matches 292; Conservative 157; Mismatches 375; Indels 307; Gaps 28;  
 QY 105 ACRNGDVERVRLVTPKVNRSRTAGKRSIPLHFAAGFGKRDVVEYLLQNGANYQARDG 164  
 Db 51 ASKEGHVKMVELLHKE-IIETTTKKGNTALHIAALAGODEVYVRELNYGANYNAQSOK 109  
 QY 165 GLIPLHNACSFHAEVYNLLLRHCGADPNARNWNTPLHEAAIKGIDVCIVLLOHGAEP 224  
 Db 110 GTPLXMAAQNENHLEVVKFLENGANQNVATEDGFTPLVALQOGHENVVAHLINYGTKG 169  
 QY 225 TI-----RNTDGRITADLADPSAKAVLTGEYKKDELLESARSNEEKMMALLPLN 275  
 Db 170 KVRLPALHIAARNDDTRTA-----AVLLQNDPNPDL--SKTG-----FTPLH 210  
 QY 276 VNCHASGRKS-----TPHLAAGYNRVKIVOLLLOHGDVHAKDKG 317  
 Db 211 IAAHYENLVNAQLLNRCASVNFPTQNGITPLHIASRRGNVIMVRLLDORGAQIETTKD 270  
 QY 318 DLVPLHNACSVGHVEVTELLVKHACVNMALWQFTPLHEAASKNRVVCVSLLSYGADP 377  
 Db 271 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDVRLLOYDAE- 329  
 QY 378 TLLNCHNKSATDLAPTQPKERLAYEFKGHSLLOAAREADVTRIKKHLSEMVNFKHPQT 437  
 Db 330 -----ID-----DIT-----LD 336  
 QY 438 HETALHCAAAAPYPRKQICELLRLKGANINEKTKFTPLHVAASEKAHNDVVEVVKHE 497



```

Db 337 HLTPLHVA--HCGHRRVAKVLLDKGAPNSRALNGFTPLHIACKKNHVRVWELLKKG 393
Qy 498 AKVNDNLGOTSLHRAAYCGHLCRLLSYGCDPNILSLOGFTALQMG-----NENVQ 552
Db 394 ASIDAVTESGLTPLHVASPMGHPVIVKLLQKQAGPNVSNVETPLHMAARAGHTEVAK 453
Qy 553 QLLOE-----GISLGNSEADRLLE-----AAKAG 577
Db 454 YLLQNKAKYNAAKDDQTPHCAARIGHTNMVKKLLLENNANPNLATTAGTTPHIAARG 513
Qy 578 DYETVKKICTVQ-SVNCRDIERQSTPLHFAAGYRNVSVEYLLHGDVHAKDKGLVP 636
Db 514 HVETVLLALEKASQACMTKG--FTPLHVAAYGKVRVAELLLERDAHPNAAAGNGLTP 571
Qy 637 LHNACSYGHEVYAEVLLHGVVNVADLWK-FTPLHEAAAKGYEICKLLHGHGADPTK 695
Db 572 LHVAHHNLDIVKLLPRGSGP-SPAWNGTTPHIAAKQVQVARSLLQYGSANAE 630
Qy 696 NRDNTPDL-LVKDGTDTQDILLRGAALLDAKGLARVKLLSPDNVNCRTDQGRHS 754
Db 631 SVQGVTPHLHAAQEGHAEMVALL-----LSQANGNLGNKSG--L 668
Qy 755 TPLHLAGAGNNLEVAEYLLHGHADYNAQDKGLIPLHNAASGHVDVAALLIKYNACVNA 814
Db 669 TPLHLVAQEGHPVADVLLKHGMVMDATRMGYTPLHVASHYGNIKLVKFLQHOADVNA 728
Qy 815 TDKWFTPLHAAQKRGRTOLCALLAHGADPTLKNQEGOTPLDL-----VS 860
Db 729 KTKLGYSPHQAQOOGHTDVTLLKNGASPNESVSDGTTPLAIKRLGYISVTDVLKV 788
Qy 861 ADVSALLTAAMPSPALPSC-----YKPO 884
Db 789 TDETSFVLVSDKHRMSFPTVDILDSDEGEELISFAERRDSRDVDEKELLDVFPK 848
Qy 885 VLVNGVRSFCATADALSSGSPSSLS-----AASLDNLGSFSELSVSVSSGTEGASS 939
Db 849 LDQVVESPA-----PRIPCAMPTVIRSEEQEAKSEYDEDSLIPSPATETSDN 900
Qy 940 LEKKEVPVGVDSITOFVNLGLEHLMDFEREQITLDVLVEMGHKELKEIGNAYGHRHK 999
Db 901 ISPVASPVHTGFLVSPM-----VDRGSGMRGSRHN 931
Qy 1000 LIKVE-----RLISGQGLNPLYTLNTSGTTLIDLSP-----1034
Db 932 GLRVVIPPRTCAAPRITRCVLKPKOKLSTPPLAEEGGLASRIIALGPTGAQFLSPVIVE 991
Qy 1035 -----DKEFOSVEEEMQSTVREHRCGHAGGIENRYNLIKIQVCN 1076
Db 992 IPHFASHGGRDELVLVLRSENGSVWKEHR-----SRYGESYLDQTLN 1033

RESULT 5
ANK2_HUMAN
ID ANK2_HUMAN STANDARD; PRT: 3924 AA.
AC Q01484; Q01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1; 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91302466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RT ankyrins reveal a family of alternatively spliced genes.";
RL J. Cell Biol. 114:241-253(1991).

[2]
RP REVISIONS.
RA Carpenter S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli E., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
RT domain and selective localization in unmyelinated axons.";
RL J. Cell Biol. 123:1463-1473(1993).
[4]
RP SEQUENCE OF 463-495 FROM N.A.
RX MEDLINE=92009921; PubMed=1833308;
RA Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
RT ankyrin gene.";
RL Genomics 10:858-866(1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA MEMBRANE OF NEURONS AS WELL AS GLIAL
CC CELLS THROUGHOUT THE BRAIN.
CC -!- PTM: PHOSPHORYLATED AT MULTIPLE SITES BY DIFFERENT PROTEIN KINASES
CC AND EACH PHOSPHORYLATION EVENT REGULATES THE PROTEIN'S STRUCTURE
CC AND FUNCTION (POTENTIAL).
CC -!- SIMILARITY: CONTAINS 23 ANK REPEATS.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X56957; CAA40278.1; -
CC EMBL; X56958; CAA40279.2; -
CC EMBL; Z26634; CAB42644.1; -
CC EMBL; M37123; AAA62828.1; -
CC PIR; S14533; S14533.
CC PIR; A39643; A39643.
CC PIR; B39643; B39643.
CC HSSP; P42771; IDC2.
CC Genew; HGNC:493; ANK2.
CC MIM; 106410; -
CC InterPro; IPR002110; ANK.
CC InterPro; IPR000488; Death.
CC InterPro; IPR000906; ZUS.
CC Pfam; PF00023; ank; 24.
CC Pfam; PF00531; death; 1.
CC Pfam; PF00791; ZUS; 1.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 21.
CC SMART; SM00005; DEATH; 1.
CC SMART; SM00218; ZUS; 1.
CC PROSITE; PS50088; ANK_REPEAT; 20.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50017; DEATH_DOMAIN; 1.
CC Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation.
FT REPEAT 63 92 ANK 1.
FT REPEAT 96 125 ANK 2.
FT REPEAT 129 158 ANK 3.
FT REPEAT 162 191 ANK 4.
FT REPEAT 193 220 ANK 5.
FT REPEAT 232 261 ANK 6.
FT REPEAT 265 294 ANK 7.

```

```
FT REPEAT 298 327 ANK 8.
FT REPEAT 331 360 ANK 9.
FT REPEAT 364 393 ANK 10.
FT REPEAT 397 426 ANK 11.
FT REPEAT 430 459 ANK 12.
FT REPEAT 463 492 ANK 13.
FT REPEAT 496 525 ANK 14.
FT REPEAT 529 558 ANK 15.
FT REPEAT 562 591 ANK 16.
FT REPEAT 595 624 ANK 17.
FT REPEAT 628 657 ANK 18.
FT REPEAT 661 690 ANK 19.
FT REPEAT 694 723 ANK 20.
FT REPEAT 727 756 ANK 21.
FT REPEAT 760 789 ANK 22.
FT REPEAT 793 822 ANK 23.
FT DOMAIN 1773 1950 REPEAT-RICH REGION.
FT REPEAT 1773 1784 REPEAT A.
FT REPEAT 1785 1796 REPEAT A.
FT REPEAT 1797 1808 REPEAT A.
FT REPEAT 1809 1820 REPEAT A.
FT REPEAT 1821 1832 REPEAT A.
FT REPEAT 1833 1844 REPEAT A.
FT REPEAT 1845 1856 REPEAT A.
FT REPEAT 1857 1867 REPEAT A.
FT REPEAT 1868 1879 REPEAT A.
FT REPEAT 1880 1891 REPEAT A.
FT REPEAT 1892 1902 REPEAT A.
FT REPEAT 1903 1914 REPEAT A.
FT REPEAT 1915 1926 REPEAT A.
FT REPEAT 1927 1938 REPEAT A.
FT REPEAT 1939 1950 REPEAT A.
FT DOMAIN 3536 3620 DEATH.
FT VARSPIC 1039 1039 Q -> QFLGKLHLTPPLNEGESLVRILQLGPCTK
      (IN ISOFORM 2).
FT VARSPIC 1444 3528 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT CONFLICT 475 476 GQ -> PE (IN REF. 4).
FT CONFLICT 971 971 I -> S (IN REF. 1).
FT CONFLICT 3581 3582 QY -> HA (IN REF. 1).
FT CONFLICT 3586 3586 I -> Y (IN REF. 1).
FT SEQUENCE 3924 AA; 430337 MW; 52AC496C428E29D2 CRC64;

Query Match 13.6%; Score 877; DB 1; Length 3924;
Best Local Similarity 29.7%; Pred. No. 2.9e-44;
Matches 293; Conservative 137; Mismatches 402; Indels 156; Gaps 23;

QY 105 ACRNGDVER-VKRLVPEKYNRSDTAGRKSTPLHFAAGFGKQDVVEYLLONGANVOARDD 163
DB 38 AARAGNLDKYVEYKGGIDINTCNQGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 95
QY 164 GGLIPLHNACSFHAEVNNLLRHGADPNARNNNYTPLEHAAIKGKIDVICVILLOHGAE 223
DB 96 KGN TALHIA LAQAEVVKVVLVEGANINAQSONGFTPLVMAQENHIDVVKYLLENGAN 155
QY 224 PTIRNTDGR-----ADLADPSAKAVLTGYKKDEL-----LESARSGNERKMMALLTPLN 275
DB 156 QSTATEDGFTPLAVALQOQHNAVAILENDTKGKVRPLPALHIAARKDDTKSAALLIQ-- 213
QY 276 VNC HASDGRKS-----TPHLAAGYNNRVKIVOLLQHGADYHAKDGLVPLPHN 324
DB 214 -NDHNADVQSKMMVNRTESGFTPLHIAAHYGNVNNATLLNLRGAADVFTARNGITPLHV 272
QY 325 ACSYGHVEVTELLVKHGACVNA MDLQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
DB 273 ASKRGNTNNWVKLLLDRGQIDAKTRDGLTPLHCAARSGHQVVELLLERGA-PLLARTKN 331
QY 385 KSAIDLAPTQPKRLAYEKGHSLLOQAAREADVTRIKKHLSEWVNFKHP-----QTHE 439
DB 332 -----GLSPLHMAAQGDHVECVKHL-----LQHKAPVDVDTLDYL 366
QY 440 TALHCAASYPKPKKQICELLKRGKANINEKTEFLTPHVASEKAHNDVVEVVKHEAK 499
DB 367 TALHVA-----HCGHYRVTKLLDKRANPNRALNGFTPLHIAACKNRIKMYELLVKYGAS 423
```

```
QY 500 VNALDNLGOTSLHRAAYCGHLOTCRLLLSYGCDPNIISLQGTALOMGN-----ENVOQL 554
DB 424 IQAITESGLTPIHVAAPFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCL 483
QY 555 LQEGI-----SLGNSEADROLLE-----AAKAGDV 579
DB 484 LRNGALVDARAREEQTPHIAASRLGKTEIVOLLQOHMAHPDAATNGYTPHLHISAREQV 543
QY 580 ETYVKKLCVTQSVNCRDIEGRQSTPLHFAAGYNNRVVVEYLLQHGADYHAKDGLVPLHN 639
DB 544 DVASVLLLEAGAAHSLATK-KGFTPLHVAARYGSLDVAKLLQRRRAADSAGKNGLTPLHV 602
QY 640 ACSYGHVEVAELLVKHGAVNVNADLWKFTPLHEAAGKYEICKLLOLQHGADPTKKNRDG 699
DB 603 AAHYDNQKVALLLLEKASPHATAKNGYTPHIAAKNQMOIASTLLNRYGAETNIVTKQG 662
QY 700 NTPDLVL-KDGDITDIDLRLGDAALDAAKGGLCLARYKLLSSPDNVNCRDTQGRHSPLH 758
DB 663 VTPHLASQEGHTDMVTLLLDKGANIHMSTKSL-----TSLH 700
QY 759 LAAGYNNLEVAEYLLQHGADYVNAQDKGLIPLHNAASYGHVDYAAALLIKYNACVNATDKW 818
DB 701 LAAQEDKVNVDILTCHKGADQDAHTKLGYPPLIVACHYGNVKMVFLLKQGANYNKTKN 760
QY 819 AFTPLHEAOKGRTQLCALLAHGADPTLKNQEGOTPLDL-----VSADDVSALLTAAM 872
DB 761 GYTPHLQAQOQGHTHIINVLLQHGAKPNATTANGNTALATAKRLGYISVVDTLKVYTEEV 820
QY 873 PPSALPSCYKQVPLNGVRSPGATADALSSGSPSSLSAASSLDNLSPSEL-----925
DB 821 TTTTITITEK-----HKLNVPTMTTEVL-----DVSDEEGDDTMTGDGGERLPEDLK 868
QY 926 -----SSVSSSGTEGASLLEKKEVPQGVDFSIQFV--RNLGLEHLMDFEREQITLDVLY 979
DB 869 ELGDDSLPSPQFIDGMNLYRLYSEGGKSDSLRFSRSHLTLSHAS--YLRDSAVMDSDV 926
QY 980 EMGHKELKEIGINAYGHRHRLKIGVERL 1007
DB 927 VIPSHQVSTLAKAEARNRYRLSNGTENL 954

RESULT 6
ANKL_MOUSE
ID ANKL_MOUSE STANDARD; PRT; 1862 AA.
AC Q02357;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin).
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Erythrocyte.
RX MEDLINE=92345717; PubMed=1386265;
RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT regulatory domain."
RL Mamm. Genome 3:281-285(1992).
CC -!- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL
CC ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4.2, TO
CC NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE
CC CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.
CC ERYTHROCYTE ANKRYNS ALSO LINK SPECTRIN (BETA CHAIN) TO THE
CC CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC PLASMA MEMBRANE.
CC -!- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).
```



RT "Prediction of the coding sequences of unidentified human genes. VII.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro."  
 RL DNA Res. 4:141-150(1997).  
 CC -!- SIMILARITY: CONTAINS AT LEAST 27 ANK REPEATS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AB002377; BAA20833.2; -  
 DR HSSP; P80144; 2MYO.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; ank; 28.  
 DR SMART; SM00248; ANK; 26.  
 DR PROSITE; PS00088; ANK\_REPEAT; 24.  
 DR PROSITE; PS0297; ANK\_REPEAT; 1.  
 KW Hypothetical protein; Repeat; ANK repeat.  
 FT NON\_TER 1  
 FT REPEAT 46 75 ANK 1.  
 FT REPEAT 79 108 ANK 2.  
 FT REPEAT 112 141 ANK 3.  
 FT REPEAT 145 174 ANK 4.  
 FT REPEAT 178 207 ANK 5.  
 FT REPEAT 211 240 ANK 6.  
 FT REPEAT 244 273 ANK 7.  
 FT REPEAT 277 307 ANK 8.  
 FT REPEAT 311 340 ANK 9.  
 FT REPEAT 344 373 ANK 10.  
 FT REPEAT 377 406 ANK 11.  
 FT REPEAT 410 439 ANK 12.  
 FT REPEAT 443 472 ANK 13.  
 FT REPEAT 476 506 ANK 14.  
 FT REPEAT 510 540 ANK 15.  
 FT REPEAT 555 584 ANK 16.  
 FT REPEAT 588 617 ANK 17.  
 FT REPEAT 622 651 ANK 18.  
 FT REPEAT 658 687 ANK 19.  
 FT REPEAT 691 720 ANK 20.  
 FT REPEAT 724 753 ANK 21.  
 FT REPEAT 761 790 ANK 22.  
 FT REPEAT 793 823 ANK 23.  
 FT REPEAT 828 857 ANK 24.  
 FT REPEAT 861 891 ANK 25.  
 FT REPEAT 895 924 ANK 26.  
 FT REPEAT 931 960 ANK 27.  
 SQ SEQUENCE 1059 AA; 113465 MW; C1F55E6CFE494770 CRC64;

Query Match 11.8%; Score 760.5; DB 1; Length 1059;  
 Best Local Similarity 27.6%; Pred. No. 4.5e-38;  
 Matches 275; Conservative 129; Mismatches 357; Indels 235; Gaps 33;

QY 102 LFEACRNGDVERKRLV-TPEKYNRSRTAGKSTPLHFAAGFGKDVVEYLLONGANVOA 160  
 Db 18 LVQAFNGDPDEVALIFPKEDVNFQD--NEKRTPLHAAAYLGDAAIEILLISGARVNA 75  
 QY 161 RDGGLPLHNAACSFHAEVYNLLLRHGDAPNARD--WNWYTPLHEAAIKGIDVCIYLLQ 219  
 Db 76 KDSKWLTPHRAVASCSEAVQVLLKHSADVNRDKNQ--TPLHIAA----- 121  
 QY 220 HGAEPTRINDGTALDAPSAKAVLTGKYEKKDELLESARSNEEKMALLTPLNVNCH 279  
 Db 122 -----ANKAV-----KCAEALVPLLSNVN 140  
 QY 280 ASDGRKSTPLHAAAYNRKIVOLLQHGADVHAKDKGLVPLHNAACSYGHYEVTELLVK 339  
 Db 141 VSDRAGTALHAAAFSGHGEVKKLLSRGANTINAFDKDRATHAAAYMGHIEVVKLLVS 200  
 QY 340 HGACVNAAMLWQFTPLHEAASKRNVVEYCSLLSYGADPTLLNCHNKSALDAPTQPKER 399

Db 201 HGAETVCKDKKSYTPLHAAASGMISVVKYLLDLGVDMPNEPNAYGNTPLVAC----- 253  
 QY 400 LAYEFKGHSLLQAAAREADVTRIKKHLSL-EMVNFKHPQTHETALHCAAAASPYPKRKOIC- 457  
 Db 254 -----YNGQDVV-----VNELDCGAIVNOKNEKGF-TPLHFAAASHT---GALCL 295  
 QY 458 ELLLRKGANINEKTEFLTPLVHASEKAHNDVVEVVVKHEAKVNALDNLGQTSLSHRAAYC 517  
 Db 296 ELLVGVGADVNMKSKDKTPLHMTALHGRFSRSQTIIQSGAVIDCEDKNGNTPLHIAARY 355  
 QY 518 GHLQTCRLLSYCGDPIIISLQGTALQMG-----NENVOQLLOEGISLGN-SEADQQLL 571  
 Db 356 GHELLINTLITSGADTAKRGIHGMFPLHIAALSGFSCCRKLLSSGFDIDTPDFDGTCL 415  
 QY 572 EAKA-GDVEYVKKL-----CTVQ-----SYNCRDI 596  
 Db 416 HAAAAGNLECLNLLNTGADFNKKDKFGSRPLHYAANANCYQCLFALVYSGASVN--DL 473  
 QY 597 EGROSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGLVPLHNAACSYGHYEAELLVKH 655  
 Db 474 DERGCTPLHYAATSDTDGKCLEYLLLRNDANPGIRDGOGYNNAVHYSAAYGHRCLQLIASE 533  
 QY 656 GAV-----VNVADLWK-----FTPLHEAAAKGYEICKLLOLQHGADPTKKNRDNTP 703  
 Db 534 TPLDVLMTSGTDMLSDSNDRATISPLHIAAYHGHQALEVLVQSLDLDRVNSSGRTP 593  
 QY 704 DLVK-DGDDTDIOLLRGDAALLDAAKKGLARVKKLSSPDNVNCRDPTQGRHSTPLHLAG 762  
 Db 594 DLAAFQGHVECDVVLINQGGASI-LVKDYILKR-----TPHAAAT 632  
 QY 763 YNNLEVAEYLLQHGADVNA---QDKGLIPLHNAASYGHVDVAALLIKYNACYNATDKWA 819  
 Db 633 NGHSECLRLLIAGNAEPQNAVYDIOGNGOTPLMLSVLNGHTDCVYSLNKGANVDAKDKWG 692  
 QY 820 FTLPLHEAAQKRTQLCALLAHAGADPTLKNOGOTPLDLVSA-----DDVSALLTAAMP 875  
 Db 693 RTALHRAVGTGHEECVDALLQHGAKCLLRDSRGRTPLHSAACGHIGVLGALLQSAASMD 752  
 QY 876 ALP-----SCYKPOVLNG-----VRSPGATADALSSGSPSS 908  
 Db 753 ANPATADNHGYTALHWACY-----NGHETVELLLEQVEFQKTEGNAFSLPHCAVINDE 807  
 QY 909 LSAASLLDNLSSGSELSVSSSGTEGASL-----EKKEYPGVDFS 951  
 Db 808 GAAEMLIDTLG-----ASIVNATDSKGRTPHAAAFTHVECLQLLLSHNAQVNSVD-- 859  
 QY 952 ITQFVRNLGLEHLMDFEREQI-TLDVLVEMGHKEL 986  
 Db 860 -----STGKTPLMMAAENGQTNVEMLVSSASAE 889

RESULT 8  
 YB23\_HUMAN  
 ID YB23\_HUMAN STANDARD; PRT; 768 AA.  
 AC Q9ULJ7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein KIAA1223 (Fragment).  
 GN KIAA1223.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,  
 RA Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XV.  
 RT The complete sequences of 100 new cDNA clones from brain which code

for large proteins in vitro.\*;

DNA Res. 6:337-345(1999).

-!- SIMILARITY: CONTAINS AT LEAST 14 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB033049; BAA86537.1; -

DR HSP; P42771; IDC2.

DR InterPro; IPR002110; ANK.

DR Pfam; PF00023; ank; 14.

DR SMART; SM00248; ANK; 13.

DR PROSITE; PS50088; ANK\_REPEAT; 13.

DR PROSITE; PS50297; ANK\_REPEAT; 1.

KW Hypothetical protein; Repeat; ANK repeat.

FT NON\_TER 1 1

FT REPEAT 1 11 ANK 1.

FT REPEAT 15 44 ANK 2.

FT REPEAT 48 82 ANK 3.

FT REPEAT 86 115 ANK 4.

FT REPEAT 119 148 ANK 5.

FT REPEAT 152 181 ANK 6.

FT REPEAT 185 214 ANK 7.

FT REPEAT 218 247 ANK 8.

FT REPEAT 251 280 ANK 9.

FT REPEAT 284 313 ANK 10.

FT REPEAT 317 346 ANK 11.

FT REPEAT 350 379 ANK 12.

FT REPEAT 383 412 ANK 13.

FT REPEAT 416 446 ANK 14.

SQ SEQUENCE 768 AA; 82819 MW; 2913B69BE2DFE06D CRC64;

Query Match

Best Local Similarity 8.8%; Score 571; DB 1; Length 768;

Matches 203; Conservative 116; Mismatches 262; Indels 256; Gaps 18;

QY 184 LLRHGADPNARDNWNYPPLHEAAIKGIDVCIVILQHGAEPIRNTDGTALDLADPSAK 243

DB 2 LLQGAENVKADNERTALIAAAGHREIVHELDHGAENVHEDVDGRTLSVA----- 56

QY 244 AVLGTGEYKKDELLESARSNEEKMALLTPLNVNCHSDGRKSTPLHLAAGYNRVKIVOL 303

DB 57 -----ALCVP-----ASKGHAS-----VVSLL 72

QY 304 LLOHGADYHAKDKGLVPLHNACSYGHYEVTETLLYKHGACVYAMDLWOTPLHEAASKNR 363

DB 73 LIDRGAENVHCDKQGMTPLLVAAAYEGHVDVOLLLEGGADVBDHTDNGRTPLLAASMGH 132

QY 364 VEVCSLLSYGADPTLLCHNKSATDLAPTLPOLKELAYEFKGSLLQ-AAREADVTRIK 422

DB 133 ASVWTLFWGA-----AVD-----SIDSGRTVLSTASAGNVVVR 170

QY 423 KHLSEMVNFKHPOTHETALHCAASPYKPKKQICELLRLKGANINEKTEPLTPHVAS 482

DB 171 TLLDRGLDENHRDDAGWTPHMAA---FEGRHICEALIEQARTNEIDNDRIFILAS 227

QY 483 EKAHNDVEVVVVKHAKVNALDNLTQTSLHRAAYCGHILQTCRLILSYGCDPNILSQFT 542

DB 228 QEGHYDCVOILLENKSNIDQRYDGRNALRVAALGHRDIVELEFSGAD----- 277

QY 543 ALQMGNNVQQLQBGISILGNSEADRLLEAKAGADVTETVKLCIVQSNCRDIEGROST 602

DB 278 -----VNCCKADGRPT- 288

QY 603 PLHFAAGYNRVSVVEYLQHGADYHAKDKGLVPLHNACSYGHYEVAELLVKHGAVNVA 662

DB 289 -LYILALENQLTMAEYFLENGANVEASDAEGRTALHVSQWGHMEMVQVLYAHADVNA 347

QY 663 DLWKFTPLHEAAAKGVEICKLLLOHGDADPTKKNRDGNTPLDLVKDGDTDIDLLRGDA 722

DB 348 DNEKRSALQSAAGWGHVVKVOLLIEHG-----A 375

QY 723 LLDAKAGKCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLIQHGADVNAQ 782

DB 376 VVD-----HFCNOGATALCIAAQEGHIDVVQVLEHGDADNHA 413

QY 783 DKGGLPLHNAASYGHVDVAALLIKYNACVNTDKWAFPLHEAAKGRQTOLCALL----- 838

DB 414 DQFGRTAMRVAANKGHSQIIKLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKVSQSLT 471

QY 839 -----LAHG-----ADPTLKNQSGOTPLDIVSADDYSAL 867

DB 472 IKSNSSCGTGGDMQPSLRGLPNGPHAFSPSPESPDSTVDRQKSSUNSLSSKSSSL 531

QY 868 LTAAMPSPA-----LPSCYKQVNLNGVRSPGATADALSSGSPSP 906

DB 532 RTTSSTATAQTVPIDSFHNLSTTEIQIQQHSLPRSRQ---SIVSPSTTQSLGQSHNSP 588

QY 907 SS-----LSAASSLDNLSSFSSELSVVSSSGTEG--ASSLEKKEVPGVDFSIQFVR 957

DB 589 SSEFEWSQVKPSLUKTKASKGKSENSAKSGSAGKAKQSNSSQPKVLEYEMTQFDR 645

RESULT 9

LATA LATMA

ID LATMA STANDARD; PRT; 1401 AA.

AC P23631; Q25328;

DT 01-NOV-1991 (Rel. 20, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Alpha-latrotoxin precursor.

OS Latrotoxin mactans (Black widow spider).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

OC Araneomorphae; Entelegynae; Araneidae; Theridiidae; Latrodectus.

OX NCBI\_TaxID=6924;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Tredecimquattus; TISSUE-Venom gland;

RX MEDLINE=91031994; PubMed=1977615;

RA Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;

RT "Cloning and structure of cDNA encoding alpha-latrotoxin from black

RT widow spider venom."

RL FEBS Lett. 270:127-131(1990).

RN [2]

RP PARTIAL SEQUENCE.

RC TISSUE-Venom;

RX MEDLINE=91362695; PubMed=1888339;

RA Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;

RT "Structure of tryptic fragments of a neurotoxin from black widow

RT spider venom."

RL Bioorg. Khim. 17:437-441(1991).

CC -!- FUNCTION: Binds to high-affinity receptors that are localized in

CC the presynaptic plasma membrane of the nerve terminal. Binding of

CC toxin results in calcium influx into nerve terminals, activation

CC of synaptic vesicle exocytosis, and ATP depletion.

CC -!- PTM: ALPHA-LATROTOXIN MAY BE PROCESSED AT ITS C-TERMINUS.

CC -!- SIMILARITY: CONTAINS 21 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X55009; CAA38753.1; -

DR PIR; S11527; S11527.

DR HSP; Q00420; LAMC.

DR InterPro; IPR002110; ANK.

DR Pfam; PF00023; ank; 20.



```
QY 132 KSTPLHFAAGFGKDVVEYL-----LQGANVQARDGGGLIPLHNAC-SFGHAEVNNLLRH 187
D 132 KSTPLHFAAGFGKDVVEYL-----LQGANVQARDGGGLIPLHNAC-SFGHAEVNNLLRH 187
Db 24 KDTPLHKAVML--PDAVERIRFVSKGADINVISDFKKTALHYAAKKTATPEVLTLYL 81
QY 188 GADPNARDNWNTPPLHEAAKIDKICVILVQHGAEPTIRNTDGTALDADPSAKAVIT 247
D 188 GADPNARDNWNTPPLHEAAKIDKICVILVQHGAEPTIRNTDGTALDADPSAKAVIT 247
Db 82 GTFNVNTDMFESTPLHYAVOENGLATKLLDLGADPNTKYMGQTPPLHCA---AMVDP 138
QY 248 GEYKDELLESARGNEKMMALLPLNVNCHASGRKSTPLHLAAGY-NRVKIYQLLQ 306
D 248 GEYKDELLESARGNEKMMALLPLNVNCHASGRKSTPLHLAAGY-NRVKIYQLLQ 306
Db 139 G-----PELVILVEYGANVNALDNKHNTPLALAEISNTWKTETLIE 182
QY 307 HGADVHAKDKGLVPLHNAC-SYGHYEVEYTELIVKHCACVNMADLWQFTPLHEAAS---K 361
D 307 HGADVHAKDKGLVPLHNAC-SYGHYEVEYTELIVKHCACVNMADLWQFTPLHEAAS---K 361
Db 183 LGADVKKINNDGITPLHLAAKSSDSKVTETLILHGADVNAATCSEGTPLHDAATSYELS 242
QY 362 NRVEVCSLLSGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKSHLLOAAREADVTRI 421
D 362 NRVEVCSLLSGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKSHLLOAAREADVTRI 421
Db 243 NTIE---MLIEYGAENAN----- 259
QY 422 KHLSEVMNFKPHQTHETALHCAASAPYKPKQICELLKRGANINEKTRKFLPLHVA 481
D 422 KHLSEVMNFKPHQTHETALHCAASAPYKPKQICELLKRGANINEKTRKFLPLHVA 481
Db 260 -----SVGDTPLHCAARSNPYHK--LKTLIAHGSNVNAVNGISVTPPLHIA 303
QY 482 ---SEKANDVVEVYKHAENALDNLGOTSILHRAAYCGHLOTCLRLLSYGCDDPNIISLQ 539
D 482 ---SEKANDVVEVYKHAENALDNLGOTSILHRAAYCGHLOTCLRLLSYGCDDPNIISLQ 539
Db 304 TYSDNA-TEALKVLTEHGAENSVDIYGRTPMH----- 335
QY 540 GFTALQMGNNVQQLLOEGISLGNSEADQLLEAAKAGADVETVKKLCITVQSVNCRDIEGR 599
D 540 GFTALQMGNNVQQLLOEGISLGNSEADQLLEAAKAGADVETVKKLCITVQSVNCRDIEGR 599
Db 336 -----YISRSYSQSLK----- 347
QY 600 QSTPLHFAAGYNRVSVVEYLQHGADVHAKDKGGLVPLHNACSYGHE--VAELLVKHGA 657
D 600 QSTPLHFAAGYNRVSVVEYLQHGADVHAKDKGGLVPLHNACSYGHE--VAELLVKHGA 657
Db 348 -----TAVELLVEHGADEAKNVIGTGPLSSACNNTEYDLRLIECFIEYGA 393
QY 658 VNVADLWKFTPLHEAAKGY-EICKLLOHGADPTKKNRDNTPD-----LVKGDGT 711
D 658 VNVADLWKFTPLHEAAKGY-EICKLLOHGADPTKKNRDNTPD-----LVKGDGT 711
Db 394 DINTRDIRDETPLYSAI---KYPEIVNLMYNSASTNITKNSNITPLESATANCIGSAEI 450
QY 712 DIQDL-----LRGDAALLDAK--KGC--LARVKKLSPPONVNCRDQGRHSTPLH 758
D 712 DIQDL-----LRGDAALLDAK--KGC--LARVKKLSPPONVNCRDQGRHSTPLH 758
Db 451 IVTQIILDAFPDINKDAIFIRNKTTIECTMLIDVKESEYDINKMRSIKFNMYGLD 510
QY 759 LAAGYNNLEVAEYLQHGADVNAQD-----KGLIPLHNA 794
D 759 LAAGYNNLEVAEYLQHGADVNAQD-----KGLIPLHNA 794
Db 511 IFIRSNINILLSVSNVEDIYLEPGFLVYGNLKRKSVYAAARKFLSLKNSISILSNIT 570
QY 795 SYGHVDVAALLIKYN 809
D 795 SYGHVDVAALLIKYN 809
Db 571 TDGYNNALPIELKYN 585
```

## RESULT 11

```
V222_FOWPV STANDARD: PRT; 747 AA.
AC Q90513;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative ankysin-repeat protein FPV222.
GN FPV222.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus."
RL J. Virol. 74:3815-3831(2000).
```

```
CC CC -!- SIMILARITY: CONTAINS 14 ANK REPEATS.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; AF198100; AAF44566.1;
DR HSP; Q00420; IAWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00023; ank; 13.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00248; ANK; 9.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 38 67 ANK 1.
FT REPEAT 103 132 ANK 2.
FT REPEAT 136 165 ANK 3.
FT REPEAT 169 198 ANK 4.
FT REPEAT 202 231 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 294 323 ANK 7.
FT REPEAT 328 357 ANK 8.
FT REPEAT 361 393 ANK 9.
FT REPEAT 397 426 ANK 10.
FT REPEAT 430 460 ANK 11.
FT REPEAT 464 493 ANK 12.
FT REPEAT 495 524 ANK 13.
FT REPEAT 529 559 ANK 14.
SQ SEQUENCE 747 AA; 85303 MW; 55F90AF2855C3D28 CRC64;
```

## Query Match

```
Best Local Similarity 6.2%; Score 399.5; DB 1; Length 747;
Matches 184; Conservative 126; Mismatches 306; Indels 193; Gaps 28;
QY 144 RKDVVEY--LQGANVQARDGGGLIPLHNAC-SFGHAEVNNLLRHGADPNARDNWNTP 201
D 16 KNDVYSVRKYLEKGINPKNKDNCNTMLYAVEHRYIDIILKLLDHGADPNYSSDHMT 75
QY 202 LHEAAIKGI-DVCITVLQHGAEPTIRNTDGTALDADPSAKAVITGKYEKKDELLESAR 260
D 76 LHSVSVIPIRKISKILTYKGMVLTANTRTFFV-----YNNRNLEIA- 120
QY 261 SGNEEKMM-----ALLTPLNVNCHASGRKSTPLHLAAGYNRVKIYQLLQHGADVHAKDK 316
D 121 -----KMLIQNGALVNMNM-----KNITPLHIASSSGSYKMLVLLHGHGANTNLT 168
QY 317 GDLVPLHNACSYGHEVTELLVKKHCACVNMADLWQFTPLHEAASKNRVEVCSLLSYGAD 376
D 169 YGETSLHYSVSSNDLNISELLIENGNTNVNANKDSITALLIAVEIMSIDLVRLLDKGAD 228
QY 377 PTLNCHNSAIDLAPTPQLKERLAYEFKSHLLOAAREADVTRIKKHLKSLVWVNFK --- 433
D 229 TNAIGL-ERFKLYVTETKQNNILKY-----LNTNNVNTVNTMINEYIASELYDMNRNS 281
QY 434 -----HPQTHETALHCAASPYKPKQICELLKRGAN--INEKTRKFLTPHLV 480
D 282 ATSKLMFRSCFEPCTVPVTLATRKGS-----KELLEILLEYCGNPDICEKTTSTA-MHY 335
QY 481 ASEKANDVVEVYKHAENALDNLGOTSILHRAAYCGHLOTCLRLLSYGCDDPNIISLQ 540
D 336 AVIRKHYEMLNLTTRYDAYTDVKDRQONTPAHYAVKLPISSECKYL----- 381
QY 541 FTALQMGNNVQQLLOEGISLGNSEADQLLEAAKAGADVETVKKLCITVQSVNCRDIEGR 600
D 382 -----KLLKLAGA-----SFNLTRKGR- 399
QY 601 STPLHFAAGYNRVSVVEYLQHGADVHAKDKGGLVPLHNACSYGHEVAELLVKHGA-V 659
```

Db 400 -PPLTACKYNTEAVKYLIESGCDTNDVMSFTPLNAVYVEREDTVKILLESQVDP 458  
 Qy 660 NVADLWKFTPLHEAAKGYEICKLILLOHGADPTKNNR-DGNTPLDLVKDCDDTDIQLLR 718  
 Db 459 NLCDDYKEVSPITQAIKRNKNKIKMLNAGIDIKIPINECYGLHMLAALHN-----KDLL- 512  
 Qy 719 GDAALLDAAKKGLARVKVLLSPDNVNCRDQTQGRHSTPL-HLAAGYNLNLEVAEYLLQHGA 777  
 Db 513 -----KWLCTISEL-----EVNGVDD---HYVPLASYVAEISDIRIMELLIEKGL 555  
 Qy 778 DYN---AODKGGILPLHNAASHGVHVAALLI-----KYNACVNATDKWAF 821  
 Db 556 DLNKVTGPDFTMTFSATSDLRKSIIDLLISQIAADEEFSGFKINKNNIOTDKVLLR 615  
 Qy 822 PLHEAAQKGRTOCALLAHAGDPTLKNOEQGPDLVLS---ADDYSALLTAAMPSPAL- 877  
 Db 616 VYHEC-----KNQVSKM-----GEIKLGDGFTMDIYKNNRSIHVNFLARYAMQLSTID 664  
 Qy 878 -----PSCYKPKOVNGVR 890  
 Db 665 LREVPYRKYLELILNPAIKRKHILNAAK 693

## RESULT 12

ANR3 HUMAN STANDARD; PRT; 832 AA.  
 AC P57078; O96KH0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat domain protein 3) (PKC-delta-interacting protein kinase).  
 GN ANKRD3 OR DIK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Fetal kidney, and Fetal lung;  
 RA Shimizu N., Kudoh J., Shibuya K.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Park H.-S., Toyoda A., Ishii K., Taudien S., Blechschmidt K., Polley A., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P., Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H., Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand B., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehtrach H., Reinhardt R., Vaspo M.-L.;  
 RA "The DNA sequence of human chromosome 21.";  
 RT Nature 405:311-319(2000).  
 RL  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 10 ANK REPEATS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB047783; BAB56136.1; -.  
 DR EMBL; AP001743; BAA95526.1; -.  
 DR HSSP; P25963; LIKN.  
 DR Genew; HGNC:496; ANKRD3.  
 DR MIM; 605706; -.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00023; ank; 10.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00248; ANK; 10.  
 DR SMART; SM00221; STYKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 9.  
 DR Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 KW ANK repeat; Alternative splicing  
 KW DOMAIN 22 286 PROTEIN KINASE.  
 FT REPEAT 485 514 ANK 1.  
 FT REPEAT 518 547 ANK 2.  
 FT REPEAT 551 580 ANK 3.  
 FT REPEAT 584 613 ANK 4.  
 FT REPEAT 617 647 ANK 5.  
 FT REPEAT 651 680 ANK 6.  
 FT REPEAT 684 713 ANK 7.  
 FT REPEAT 717 746 ANK 8.  
 FT REPEAT 750 780 ANK 9.  
 FT REPEAT 782 811 ANK 10.  
 FT NP\_BIND 28 36 ATP (BY SIMILARITY).  
 FT BINDING 51 51 ATP (BY SIMILARITY).  
 FT ACT\_SITE 143 143 BY SIMILARITY.  
 FT VARSPIC 278 325 MISSING (IN ISOFORM 2).  
 FT CONFLICT 714 714 M -> V (IN REF. 1).  
 SQ SEQUENCE 832 AA; 91610 MW; 5D8FFD5F04F7ECB CRC64;  
 Query Match 6.0%; Score 390.5; DB 1; Length 832;  
 Best Local Similarity 28.1%; Pred. No. 6.2e-16;  
 Matches 141; Conservative 58; Mismatches 176; Indels 127; Gaps 13;  
 Qy 368 SLLLSYGADPTLLNCHNKSADLAPTPQLKERLAYEFKGHSLLQAAAREADVTRIKHLSL 427  
 Db 434 SLSLSFEREP-----STSLGTIDVQKK-----LYDAIVSGDTSKLMKTL-- 474  
 Qy 428 EMVNEKHPQTHETALHCAAA----SPYPRKQICELLRRKANINERKTEFLPLHVASE 483  
 Db 475 -----QPQDVLALDSGASLLHLAVEAGQECAKWLLNNANPNLSNRGSGTPLHMAVE 528  
 Qy 484 KAHNDVVEVVKHEAKVNALDNLGQTSLHRAACVGHLOTCTRLLSYGCDDNIIISLOGFT- 542  
 Db 529 RVRGVWELLARKTSVNAKDEQDQWTLHPAQAQNGDESSTRLLEKNASVNEVDFEGRTP 588  
 Qy 543 ---ALQMGNEN-VQOLLQEGISLGNSEADQRLLEAAKAGDVETVKLCTVQSVNCRDIEG 598  
 Db 589 MHVACQHGGENIVRIILRRGVV-----SIQG 615  
 Qy 599 ROS-TPLHFAAGYNRVSVVEYLL-QHGADYHAKDKGLVPLHNACSYGHYEVAELLVKGH 656  
 Db 616 KDALPLHYAAQGHLPVVKLLAKQPCVSNQAQTLGDRTPLHLAAQGRHYRVARILDL 675  
 Qy 657 AVNVADLWKFTPLHEAAKGYEICKLILLOHGADPTKNNRDNTPDLVKDGDGTDIQL 716  
 Db 676 SDNVNCSLLAQTPHVAETGHTSTARLLHARGAGKAMTSDGYTAUHL----- 724  
 Qy 717 LRGAALLDAAKGKGLARVKVLLSPDNVNCRDQTQGRHSTPLHLAAGYNLNLEVAEYLLQHG 776  
 Db 725 -----AARNGLHATVK-----LLVEEK 741





```
DR Pfam; PF00023; ank: 10.
DR SMART; SM00248; ANK: 6.
DR PROSITE; PS50088; ANK_REPEAT; 7.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 40 69
FT REPEAT 144 173
FT REPEAT 177 206
FT REPEAT 210 239
FT REPEAT 272 302
FT REPEAT 306 336
FT REPEAT 340 370
FT REPEAT 374 403
FT REPEAT 407 437
FT REPEAT 441 471
FT REPEAT 473 502
FT REPEAT 571 602
SQ SEQUENCE 668 AA; 76200 MW; F3C5BA076023791B CRC64;

Query Match 5.9%; Score 380.5; DB 1; Length 668;
Best Local Similarity 27.3%; Pred. No. 1.8e-15;
Matches 153; Conservative 95; Mismatches 192; Indels 121; Gaps 26;

QY 326 CSYGHVEVTELV-----KHGA--CVNMDIMQPTPLHEAASKNRVEVCSLLSYGADPT 378
   | : | : | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 11 CRGNYEILSAITNYNLHKHCANOCEN--ESIPFTAIHQALQRLQIDIVKELIQ--QNP 66

QY 379 LL-----NCHNKAIDIAPTPOLKERLAYEPKGHSLLOAAREADYTRIKKHLSELMVNF 434
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 67 LIYVTHRRNSTLHTICITINMVD-----IVISLTVDCDILDIKIYAST-ILN-KH 115

QY 435 POTHETALHCAAAASPYPKRRQICELLRLKRGANINE----KTKEFLTPLHVAASEKAHND-- 488
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 116 -----KIGEACTHVLKEGISGNEISVINKINKSIEMKLIKERTQQDEL 158

QY 489 -VVEVVVVKHEAKVNALDNLGQTSLHRAAYCYGHLQTCRLLLSYGCDPNIIISLOGFTALQ-- 545
   | : | : | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 159 LIAEMLLKGGIDVNAKDVCYRTPHYAAGERNTKMNLLLSYGADVNIITLDDLSVLEYA 218

QY 546 MGENVVOOLQEGISIGN--SEADROLLEAAKAGDVEVTKKLC--VQSVNCRDIEGRQSTP 603
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 219 VDSKNIDITIKAITDNRSNINKNDLSLKAIKNTDLETSLLYDSGFSVNSIDV--YKNTP 276

QY 604 LHFAAGYNRVS-VVEYLLQHGADVHAKDGGVPLHFNACSYGH-YEVAELVKKHGAVVNV 661
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 277 LHVAQAPSLSRVLPKLLERIGIDVNAKNIKGETPLYLMAKNKYDTNIRTLIMRGADVNA 336

QY 662 ADLWKFPTLHEAAKGYEICKL-LLOHGADPTKKNRDGNTPLDLVKDGDDTDIQDLLRGD 720
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 337 ADSLYITPLHQASTLDYKDTVTITLLELGA----- 366

QY 721 AALLDAAKGGCLARVKKLSSPDVNCRDTOGRHSTPLHLAAGYNLLEVAEYLLQHGADV 780
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 367 -----NVNARDYCDK--TPHYAAVRNVVLIINTLLDYGADIE 402

QY 781 A-QDKGGLIPLHNAASYG---HVDVAALLIKYNACVNATDKWAFPTPLHEAAQKG-RTQLC 835
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 403 ALSQKIGTV-LH-PALYGTNPYMSVKT-LIDRGANVSNKNKYLSTPLHYACKKNCKPEVI 459

QY 836 ALLAHGADPTLKNQEGOTPL 856
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 460 KMLLDGADVNAINRNQYPL 480

RESULT 15
ID ASB2_HUMAN STANDARD; PRT; 587 AA.
AC Q96Q27; Q9NSU5; Q9Y567;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ankyrin repeat and SOCS box containing protein 2 (ASB-2).
GN ASB2.
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21450655; PubMed=11566180;
RA Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,
RA Nakanishi T., Tanaka K.;
RT "ATRA-regulated Asb-2 gene induced in differentiation of HL-60
RL leukemia cells.";
RL FEBS Lett. 505:223-228(2001).
RN [2]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21634875; PubMed=11682484;
RA Guibal F.C., Moog-Lutz C., Smolewski P., Di Giola Y.,
RA Darzynkiewicz Z., Lutz P.G., Cayre Y.E.;
RT "ASB-2 inhibits growth and promotes commitment in myeloid leukemia
RL cells.";
RL J. Biol. Chem. 277:218-224(2002).
RN [3]
RP SEQUENCE OF 31-587 FROM N.A.
RC TISSUE=Uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 42-587 FROM N.A.
RX MEDLINE=20564172; PubMed=11111040;
RA Kile B.T., Viney E.M., Willison T.A., Brodnicki T.C., Cancilla M.R.,
RA Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA Alexander W.S.;
RT "Cloning and characterization of the genes encoding the ankyrin repeat
RT and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL Gene 258:31-41(2000).
CC !- FUNCTION: Not known. Its expression induces growth inhibition and
CC chromatin condensation.
CC !- INDUCTION: By all-trans retinoic acid (ATRA).
CC !- SIMILARITY: CONTAINS 12 ANK REPEATS.
CC !- SIMILARITY: CONTAINS 1 SOCS BOX DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB056723; BAB64532.1; -.
DR EMBL; AJ251238; CAC17765.1; -.
DR EMBL; AL137735; CAB70899.1; -.
DR EMBL; AF159164; AAD45345.1; -.
DR Genew; HGNC:16012; ASB2.
DR MIM; 605759; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001496; SOCS.
DR Pfam; PF00023; ank; 11.
DR PROSITE; PS50088; ANK_REPEAT; 9.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50225; SOCS; 1.
KW ANK repeat; Repeat.
FT REPEAT 56 85
FT REPEAT 89 119
FT REPEAT 123 152
FT REPEAT 156 185
FT REPEAT 189 218
FT REPEAT 222 251
FT REPEAT 255 284
FT REPEAT 288 317
FT REPEAT 320 349
FT REPEAT 362 391
FT REPEAT 392 421
```

Search completed: February 12, 2003, 03:33:06  
Job time : 20.3675 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2003, 03:26:09 ; Search time 46.6325 seconds  
(without alignments)  
5478.980 Million cell updates/sec

Title: US-09-843-159B-4  
Perfect score: 6464  
Sequence: 1 RCSARRGAAGGQAGRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4005	62.0	1181	5 Q9VBP3	Q9vbp3 drosophila
2	3997	61.8	1181	5 Q9XZ37	Q9xz37 drosophila
3	904	14.0	1136	6 Q9N180	Q9n180 bos taurus
4	893	13.8	1726	11 Q8VC68	Q8vc68 mus musculus
5	892.5	13.8	4377	4 Q12955	Q12955 homo sapien
6	891	13.8	1943	11 Q61307	Q61307 mus musculus
7	888.5	13.7	1762	11 Q88521	Q88521 rattus norv
8	885.5	13.7	2622	11 Q70511	Q70511 rattus norv
9	880	13.6	1856	4 Q99407	Q99407 homo sapien
10	878.5	13.6	1719	4 Q13768	Q13768 homo sapien
11	876	13.6	1848	11 Q61302	Q61302 mus musculus
12	868.5	13.4	1549	5 Q9V4B1	Q9v4b1 drosophila
13	866.5	13.4	1549	5 Q24241	Q24241 drosophila
14	858	13.3	843	11 P97582	P97582 rattus norv
15	858	13.3	1809	5 Q17487	Q17487 caenorhabdi
16	858	13.3	1867	5 Q17486	Q17486 caenorhabdi

17	858	13.3	2039	5	Q17489	Q17489 caenorhabdi
18	854	13.2	1815	5	Q17488	Q17488 caenorhabdi
19	854	13.2	6994	5	Q17343	Q17343 caenorhabdi
20	851	13.2	1786	5	Q17344	Q17344 caenorhabdi
21	848.5	13.1	1159	5	Q9NCP8	Q9ncp8 drosophila
22	827.5	12.8	2443	5	Q9VSA2	Q9vsa2 drosophila
23	817	12.6	1009	5	Q8SWY2	Q8swy2 drosophila
24	657	10.2	791	4	Q9UFA4	Q9ufa4 homo sapien
25	635.5	9.8	1619	5	Q9NGR8	Q9ngr8 drosophila
26	629.5	9.7	1088	4	Q13484	Q13484 homo sapien
27	627	9.7	1913	5	Q9GRV5	Q9grv5 caenorhabdi
28	609	9.4	1453	5	Q9VMR4	Q9vmr4 drosophila
29	590	9.1	1599	11	Q9NHH0	Q9nhh0 mus musculus
30	563.5	8.7	934	16	Q83807	Q83807 treponema p
31	557	8.6	1486	4	Q8TEF1	Q8tef1 homo sapien
32	554.5	8.6	1098	11	Q61304	Q61304 mus musculus
33	552	8.5	1411	5	Q02989	Q02989 latrodectus
34	539.5	8.3	1395	5	Q9XZC0	Q9xzc0 latrodectus
35	539	8.3	833	4	Q96186	Q96186 homo sapien
36	530.5	8.2	1282	5	Q8T4F8	Q8t4f8 drosophila
37	530.5	8.2	2119	5	Q9VAU5	Q9vau5 drosophila
38	530.5	8.2	2119	5	Q8T9G9	Q8t9g9 drosophila
39	527	8.2	743	4	Q8TB46	Q8tb46 homo sapien
40	525.5	8.1	4001	5	Q8WRQ7	Q8wrq7 drosophila
41	524	8.1	1188	4	Q9H288	Q9h288 homo sapien
42	509	7.9	1021	13	Q8UVC1	Q8uvc1 brachydantio
43	507.5	7.9	1058	16	Q9PCW4	Q9pcw4 xylella fas
44	505.5	7.8	686	4	Q9NXF0	Q9nxf0 homo sapien
45	495	7.7	861	5	Q8WP35	Q8wp35 geodia cydo

## ALIGNMENTS

## RESULT 1

Q9VBP3	PRELIMINARY;	PRT;	1181 AA.
ID Q9VBP3			
AC Q9VBP3			
DC Q9VBP3			
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE BCDA:LD22548 protein.			
GN BCDA:LD22548 OR CG4719 OR CG17487.			
OS Drosophila melanogaster (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;			
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephyroidae; Drosophilidae; Drosophila.			
OX NCBI_TaxID=7227;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=BERKELEY;			
RX MEDLINE=20196006; PubMed=10731132;			
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,			
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne B.D.,			
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
Jalali M., Kalush F., Karpis G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milchina N.V., Mobarly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003752; AAF56487.1; -.  
DR HSP: Q00420; IAWC.  
DR FlyBase: FBgn0027508; BcDNA:LD22548.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR001660; SAM.  
DR Pfam: PF00023; ank; 17.  
DR Pfam: PF00536; SAM; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 15.  
DR SMART: SM00454; SAM; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 14.  
DR PROSITE: PS50297; ANK\_REPEAT; 1.  
KW ANK repeat; Repeat.  
KW ANK repeat; Repeat.  
SQ SEQUENCE 1181 AA; 127944 MW; EB6F0F94E6A3D504 CRC64;

Query Match 62.0%; Score 4005; DB 5; Length 1181;  
Best Local Similarity 65.3%; Pred. No. 7.4e-242;  
Matches 772; Conservative 166; Mismatches 189; Indels 56; Gaps 13;

QY 91 ABAVPEPAARELFACRNGDVERVKRLVTEPKVNSRDTAGRKSTPLHFAAGFGKDVVEY 150  
DB 17 AVMANDP-LRELEFACTGEGIAKVKLITPOTNARDTAGRKSTPLHFAAGFGREVEF 75  
QY 151 LILONGANVOARDGGLPLHNACSFHAEVYVLLLRHAGDPNARDNNYTPLEAAIKGK 210  
DB 76 LINSAGSIQACDEGLPLHNCSFGHAEVYVLLLRHAGDPNARDNNYTPLEAAIKGK 135  
QY 211 IDVCIVLHOGAEPITINTDGRALDADPSAKAVLFGYKKDELLSAGSGNEKMMAL 270  
DB 136 VDVCALLQHGANTIRNSQKTPLELADEATREVLTVGEYRKDELLSAGSGNEKMMAL 195  
QY 271 LTPLVNCHASDGRKSTPLHLAGYVNVKTVOLLHOGADVHAKDKGLVPLHNACSYGH 330  
DB 196 LTPLVNCHASDGRKSTPLHLAGYVNVKTVOLLHOGADVHAKDKGLVPLHNACSYGH 255  
QY 331 YEVTLLVHKGACVAMDLWQFTPLHPEAASKRVEVCSLLLSYGADPTLLNCHNKSIDL 390  
DB 256 FDTYKLLIQGANVANDLWAFPLHPEAASKRVEVCSLLLSRGADPTLLNCHNKSALDA 315  
QY 391 APTPOLKERDAYEFKSHLSLQAAAREADVTRIKKHLISLEVMFKHPQTHETALHCAASPY 450  
DB 316 APTRELRETAPEYKGCILLDACRCKDVSRAKLVCAEIVNFVHPYTGDTPLHLAVVSPD 375  
QY 451 PKRQICELLRKCANINTEKTEFLTPLVHASEKAHNDVVEVVKHAKVNALNLSGTS 510  
DB 376 GKRQKMLLETRKGLNLNENKNAFLTPHLAAEILLHYDAMEVLLKQKAGNALDSLOQTP 435  
QY 511 LHRAAYCGHLQTCRLLSYCDPNIIISLQGFALQMGNEVNVQQLQEGISLGNSEADROL 570  
DB 436 LHRCA--RDQAVRLLSYAADTNVISLEGLTAQLASDSVKLLK-----NPPDSETHL 488  
QY 571 LEAAKAGDVETVKKLCITVQ--SYNCRDIEGRQSTPLHFAAGYVNSVVEYLLQHCADVHA 628  
DB 489 LEAAKAGDLDTVRRILNNPISVNCRLDRGHSHTPLHFAAGFNRPVVPVQFLEHGAEYVA 548  
QY 629 KDKGGLVPLHNACSYGHVEAEVLLVKGAVNVNADLWKFPTPLHPEAASKGVEICKLILQH 688

DB 549 ADKGLVPLHNACSYGHVEAEVLLVKGAVNVNADLWKFPTPLHPEAASKGVEICKLILQH 608  
QY 689 GADPTKKNRNDGTPUDLVKDGDTIDQDLLRGDAALLDAAKGCLARVKKLSSPDNVCNRD 748  
DB 609 GADPMKKNRNDGATPADLVKESDHDVAELLRGPSSALLDAAKGNLARVORLVPESINCRD 668  
QY 749 TQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKY 808  
DB 669 AQGRNSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKH 728  
QY 809 NACVNATDKWAFPTPLHPEAASKRTOICALLLAHAGDPTLKNQGGPPLDVSADDSALL 868  
DB 729 KTVVNATDKWGFPTPLHPEAASKRTOICALLLAHAGDAYMKNGGQPIELATADDVKCLL 788  
QY 869 TAAMPSPS---ALPSCYKQVPLNGVRSGATADALSSGSPSSPSS--LSAASSLDNLSSGF 922  
DB 789 QDAMATSLSQOAL-SASTQSLTSSSPADPATAAAGPTSSSSSAILSPTTETVLLPTGA 847  
QY 923 SELSSV-----VSSS-----GPEGASSLEKKEVPGVD--FSITQFVFNRLGLEH 963  
DB 848 SMILSVVPVPLPLSSSTRISPAQGAENGAGSSSDDL--LPDADTTINVSFGFLSSQOLHH 905  
QY 964 LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLKGVRLISSQOGLNPLYTLNTS 1023  
DB 906 LIEUFEREQITLDILAEMGHDDDLKQGVSAFGFRHKLKGLIAQLRS-----TT 953  
QY 1024 CSG-----TILIDLSPDDKFEQSVVEEMQSTVREHRDGGHAGGIFNRYNLIKQKVCNK 1077  
DB 954 GIGNNVNLTLLVOLLPPDKFEVVEEMQATIREHRDNGOAGGYTRYNIIRVQKQVR 1013  
QY 1078 KLWEYTHRRKEVSEENHNANERMLFHGSPFVNAILHKGDFDERHAYIGMFGAGYFAE 1137  
DB 1014 KLWEYTHRRKEVSEENHNANERMLFHGSPFVNAILHKGDFDERHAYIGMFGAGYFAE 1073  
QY 1138 NSSKNQVYVGGTGGTGVPHKDRSCYICHRQLFCRVLTKGKFLQFSAMKMAHSPGHH 1197  
DB 1074 HSSKNQVYVGGTGGTGGTGVPHKDRSCYICHRQLFCRVLTKGKFLQFSAMKMAHSPGHH 1133  
QY 1198 SVTGRSYNGVLALEYVIRGEQAYPEYLITYQIMPEGMVDG 1240  
DB 1134 SVGRPSAGSLHFAEYVYVYRGQSYPEYLITYQIVKPDSSSG 1176

RESULT 2  
Q9XZ37  
ID Q9XZ37 PRELIMINARY; PRT: 1181 AA.  
AC Q9XZ37;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 127.9 kDa protein.  
GN BCDNA:LD22548 OR CG4719 OR CG17487.  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OC NCBI\_TaxID=7227;  
RP [1]  
RP SEQUENCE FROM N.A.  
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,  
RA Agbayani A., Arcina T.T., Baxter E., Blazej R.G., Butenhoff C.,  
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,  
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,  
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,  
RA Park S., Sequeira A., Sethi H., Snir E., Svirska R.R., Weinburg T.,  
RA Gelniker S.E.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF132196; AAD34784.1; -.  
DR HSP: Q00420; IAWC.  
DR FlyBase: FBgn0027508; BcDNA:LD22548.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR001660; SAM.



Db 211 IAAHYENLVAQILLNKGASVNFPTQNGIITPLHFIASRRGNVIMVRLLLDRGAQIETRTKD 270  
QY 318 DLVPLHNACSYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCSSLISYGA-- 375  
Db 271 ELTPLHCAARNGHURISEIILLDHGAPTOAKTKNGLSPIHMAAQGDHDDCVRELLQYNAEI 330  
QY 376 -DPTL-----LNC-HNKA---IDLAPTQKRLAYEFKGSILLOAREADVTIRI 421  
Db 331 DDITDLHLTPHVAACHGHHRAKVLDDKGAKPNR-----ALNGFTPLHTACKKNHTRV 385  
QY 422 -----KKHLSLEWNVKHPOTHETALHCAASPYPKKQICELLRKAGANINEKTEP 476  
Db 386 MELLKMGASIDAVT-----ESGLTPLHVASFMGHP---PIVKSLLQREASPNVNVK 438  
QY 477 PLHVAEKANDVVEVVVKEAKVNALDNLGOTSLHRAAYCGHLQTCRLLISYGCDPNII 536  
Db 439 PLHMAARAGTEVAKYLDQWAKVNAKAKDDOTPLHCAARIGHTNMVKLLLENANPNLA 498  
QY 537 SLOGFTALONGENVQOLLOEGISLGNSEADRLLEAAKAGDVETVKKICTVQ--SVNCRD 595  
Db 499 TTAGHTPLHI-----AAREGHVETALALEKEASQTCMT 532  
QY 596 IEGROSTPLHFAAGYNRVSVVYLLQHGADVHAKDKGGLVPLHNACSYGHEVVAELLVKH 655  
Db 533 KKG--FTPLHVAAYKYGKVRMAELLEDHDAHPNAAGKSGLTPLHVAVHHNHLDVVRLLPR 590  
QY 656 GAVVNVADLWKFTPLHEAAAKGYEICKLLLOHGADPTKKNRDGTPLDL--VKDGDTDIQ 714  
Db 591 GGSFHPALNGYTPHIAAKONQLEVARSLLOYGGSNAESVOGYTPHLHAAQEGHAEV 650  
QY 715 DLLRGDAALLDAKAGCLARKVKKLSSPDNVNCRDQTQGRHSTPLHLAGYNLEVAEYLLQ 774  
Db 651 ALL-----LSKANGNLGNKSG--LTPHLHVAQEGHIPVADVLIK 688  
QY 775 HGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWFTPLHAAQKGRQL 834  
Db 689 HGVTVDATRMGYTPPLHVAHYGNIKLVKFLLOHKADVNAKTKILGYSPLHQAQOQHPTDI 748  
QY 835 CALLAHGADPTLKNBQGTPLDL-----VSADDVSALITTAAMPSPALPCYK---PQV 885  
Db 749 VTLHLKHGASPNVSSNGTTPLAIAKRLGYSIVTDVLKVVYDDETTVQLITDKHRMSPET 808  
QY 886 LN-----GVRSPGATA-----DALSSGSPSP-----SS 908  
Db 809 VDEILDVSEDEGDLGPRABRPARDVDEKELLDVFPKLDQVVESPAIPRIPCVTPTV 868  
QY 909 LSAASSLDNLGSGFSELSSVSVSSGTGASLEKKEKPGVDFTSQFVRNLGLEHLMDF 968  
Db 869 VIRAEPDQVSKYEDESLIPSPATETSDNISPVASPVHTGFLVSF----- 916  
QY 969 FREQITLDVLVEMGHKELKEIGINAYGHRHKLKIGVE-----RLISGQOGLN 1015  
Db 917 -----VDAGGSMRSGRHNGLRVVPPRPTCAAPRITCRILVRPKLPT 959  
QY 1016 PYLTLNLTSGGTILIDLP-----DDKEFQSVSEEMOSTVVRHR 1054  
Db 960 PPLAAEEGLASRVIALGPTGVQFLSPVIVEIPHASQGRGDRVLRSENGSVWKEHK 1019  
QY 1055 DGHGAGGIFNRYNLTKQVCN 1076  
Db 1020 -----NRYGESYLDQILN 1032  
RESULT 4  
ID Q8VC68 PRELIMINARY; PRT; 1726 AA.  
AC Q8VC68;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 188.2 kDa protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary Gland;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
EMBL: BC021657; AAH21657.1; --  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000488; Death.  
DR InterPro: IPR000906; ZUS.  
DR Pfam: PF00023; ank; 24.  
DR Pfam: PF00531; death; 1.  
DR Pfam: PF00791; ZUS; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 22.  
DR SMART: SM00005; DEATH; 1.  
DR SMART: SM00218; ZUS; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 21.  
DR PROSITE: PS50297; ANK\_REPEAT; 1.  
DR PROSITE: PS50017; DEATH\_DOMAIN; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 1726 AA; 188239 MW; 430148B67E00B1CA CRC64;  
Query Match 13.8%; Score 893; DB 11; Length 1726;  
Best Local Similarity 28.3%; Pred. No. 8.1e-47;  
Matches 337; Conservative 167; Mismatches 497; Indels 188; Gaps 38;  
QY 68 AVAAARIMSGRCAGGGAACASAAAEVPAARELFEACRNGDVERKRLVTEPKVNSRD 127  
Db 10 AKPAIRKRRKKKSDANASYLRAARAGHLEKA---LDYIKNG-----VDVNICN 54  
QY 128 TAGRSTPLHFAAGFGKRDVWEYLLQNGANVOARDGGLIPLHNACSFHAEVYNLLRH 187  
Db 55 QNGLNA--LHLASKEGHVEVSELLQREANVDAATKKGNTALHIASLAGQEVKVLVTN 112  
QY 188 GADPNARDNNWYTPHLEAAIKGIDYICIVLLOHGAETPIRNTDGTALDLA-----DPSA 242  
Db 113 GANVAQSQNGTPLYMAAQENHLEVVRFLDNGASQSLATEDGFTPLAVALQGGHQQV 172  
QY 243 KAVLTGEYK--KDELESARSGNEEKMWALLTPLNVNCHASDGRKSTPLHLAGYNRVK 299  
Db 173 SULLENDTKGVKPLPALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNIN 232  
QY 300 IVQLLQHGADVHAKDKGLVPLHNACSYGHEVTELLVKGACVNAAMDLMWQFTPLHEAA 359  
Db 233 VATLLNRAAAVDFTARNDIPLHVASKRGKGNANWVLLDRGAKIDAKTRDGLTPLHCGA 292  
QY 360 SKNRVEVCSSLISYGADPTLLNCHNKSAIDLAPTQKRLAYEFKGSILLOAREADVT 419  
Db 293 RSGHEQVVEMLDRSA--PILSKTN-----GLSPLHMATQGD-- 328  
QY 420 RIKKHLSLEWNVFKH-----PQTHETALHCAASPYPKKQICELLRKAGANINEKTK 472  
Db 329 ----HLNCVOLLQHNVPVDDVTNDYLLALHVAACHCHYK---VAKVLLDKKASPNKAL 381  
QY 473 EFLTPLHVAEKANDVVEVVVKEAKVNALDNLGOTSLHRAAYCGHLQTCRLLISYGD 532  
Db 382 NGFTPLHIAACKNRIRVMELLKHGASIQAVTESGLTPIHVAAFMGHVNIVSOLMHGAS 441  
QY 533 PNLISLOGFTALQM-----NENVQOLLQEGISLGNSEADRO--LLEAAKAGDVETVKKL 585  
Db 442 PNTTVNRGETALHMAARSGAEVRYLVODGAQVAKAKDDQTPHLISALRGADIVQOL 501  
QY 586 CTV-OSVNCRDIEGROSTPLHFAAGYNRVSVVYLLQHGADVHAKDKGGLVPLHNACSYG 644  
Db 502 LOQGASPNAAITSG--YTPHLHAAAREGHEDVAAFLDHGASLSITTKGFTPLHVAAYK 559  
QY 645 HYEVAELVKGAVNVNADLWKFTPLHEAAAKGYEICKLLLOHGADPTKKNRDGTPLD 704  
Db 560 KLEVASLLQKSPADPAACKSGLTPLHVAHYDNQKVALLLDQGAASPHAAAKNGYTPLH 619  
QY 705 L-VKGDGTDIDILL-----RGDAALLDAKKGCLARVKKLSSPD-VNVCRTQ 750



Db 620 IAAKNOMDIATSLLEYGADANAVTRQGIASVHAAQEGHVDVMSLLSRNANVNLNKS 679  
QY 751 GRHSTPLHLAAGYNNLEVAEYLLQHGADYNAQDKGLIPLHNAASGYCHVDVAALLIKYNA 810  
Db 680 G--LPLHLAAQEDRVNAEVLVNOGAHVDAQTKMGYTPPHVCHGYGNIKIVNFFLQHS 737  
QY 811 CVNATDKWFTPLHEAAQKRTOLCALLAHGADPTLKNOEGQTPLDLVS-----ADD 863  
Db 738 KVNAKTKNGYTALHQAQOQTHIINVLLQNNASPNELTVNGNTALAIARRLGIYSVVD 797  
QY 864 VSALLTAAMPSPALPCYK---POVLNGV--RSPGATADALS-----GP----- 903  
Db 798 LKVTTEEIMTTTITEKHKMNPETMNEVLDMSDDGEADITGDTDKYLGPDQDKELGDD 857  
QY 904 SSPSLSAASLDNLSGSPSELSSVVS-----SSGTEGASSLEKKEKVPVGVDFSIIOFVN 958  
Db 858 SUPAEGYGVFSLGARSASLSFSSDSRYTLNRSSYARDSMMTEELLVPSKEQHLT-FTRE 916  
QY 959 LGLEHLMDFEREQITLDVL-----VEMGHKELKEIGINAYG-----HRHKLII--- 1001  
Db 917 FDSDSLRLH--YSWAADTLDNVLNVSSPVHSGF--LVSMFVDARGGSMRGRHGMRIIIP 973  
QY 1002 -----KGVRLISQOGLNPYLTLNTSGSGTILIDLSPDKETQS-VEEEMQSTVREH 1053  
Db 974 RKTAPTTRITCLVRKHLKLANPPMVVEGEGLASRLVEMGPAGAFGLGPVIVEIP----- 1027  
QY 1054 RDGHAGGIFRNYLNIKQKCNKLWE--RYTHRKVEYSE-----ENHNHANERMLPHG 1106  
Db 1028 ---HFGSMRGRERELIVLRSENGETWREHOFDSKNEDLAELLNGMDELDSPPEL---G 1080  
QY 1107 SPFVNAITHKGFDERHAYIGMGFGAGIYFAENS---SKSNQYVYIGGG 1152  
Db 1081 TKRICRIITKDPQ-----YFVAVSRIKQESNQ--IGPEGG 1114

RESULT 5

Q12955 PRELIMINARY; PRT; 4377 AA.  
ID Q12955  
AC Q12955;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Ankyrin G.  
GN ANK-3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRIN STEM;  
RX MEDLINE=95138209; PubMed=7836469;  
RA Kordeli E., Lambert S., Bennett V.;  
RT "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized  
at the axonal initial segment and node of Ranvier.";  
RT J. Biol. Chem. 270:2352-2359(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRIN STEM;  
RA Carpenter S.S.;  
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U13616; AAA64834.1; -;  
DR HSSP; P55273; 1B18.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00023; ank; 24.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 21.  
DR SMART; SM00005; DEATH; 1.

DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 21.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
KW ANK repeat; Repeat.  
SQ SEQUENCE 4377 AA; 480395 MW; F42379E55768B684 CRC64;  
  
Query Match 13.8%; Score 892.5; DB 4; Length 4377;  
Best Local Similarity 30.7%; Pred. No. 3.7e-46;  
Matches 277; Conservative 126; Mismatches 365; Indels 135; Gaps 21;  
  
QY 136 LHFAAGRGKRDVVEYLLONGANYQARDGGLIPLHNACSFGEAEVNNLLRHGHGADPNARD 195  
Db 78 LHLASKEGHEVVEYSELLQREANVDAATKKNTALHTASLAGQAQAEVVKVLTNGANVNAQS 137  
QY 196 NNNYTPPLHEAAIKGKTDVCLVILLOHGAETIRNTDGTALDLA-----DPSKAVLTGEY 250  
Db 138 QNGFTPLYMAAQENHLEEVVKFLDNGASOSLATEDEGFTPLAVALQOQHDOVVSLLLENDT 197  
QY 251 K---KDELLESARGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLQ 307  
Db 198 KGKVRPLPALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINATLLNR 257  
QY 308 GADVHAADKGDVPLHNACSYGHEVTELLYKHGACVNAAMLQFTPLHEAASKNRVEVC 367  
Db 258 AAADVFTARNITPLHVASKRGNNAMVKLLDRGAKIDAKTRDGLTPLHCGARGHEQVV 317  
QY 368 SLLSYGADPTLLNCNKSNAIDLAPTQPKERLAYEFKHSLLQAAEADVTIRIKHLSL 427  
Db 318 EMLLDRAA--PILSKTKN-----GLSLPHMATQGD-----HLNC 349  
QY 428 EMVNFKH-----PQTHETALHCAASPYPRKROICEILLRRKGANINEKTEFLPLHV 480  
Db 350 VQLLLQHNVPDVTNDYTLALHVAACHGK---VAKVLLDKKANKAKALNGFTPLHI 406  
QY 481 ASEKANDVVEVVKHEAKVNALDNLGOTSILHRAAYCGHILQTCRLLLSYGCDPNIIISLQ 540  
Db 407 ACKKNRIKVNELLKHGASQAVTESGLTPIHVAEFMGHVNIVSOLMHHGASPNNTNVRG 466  
QY 541 FTALQMG-----NENVOQLLOEGISLGNSEADRO---LLEAAKAGDVETVKKLCIV--QSVN 592  
Db 467 ETALHMAARSQAQAEVVRVYLVQDGAQVEAKAKDDOTPLHISARLGKADI VOQLLQOQASPN 526  
QY 593 CRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHEVAELL 652  
Db 527 AATTSG--YTPLHLSAREGHEDVAFLDDHGASLSITTKGFTPLHVAAYKGLVANLL 584  
QY 653 VKHGAVVNADLWKFTPLHEAAAKGYETCKLLQLQHGADPTKKNRDGNTPLDL-VKDGDT 711  
Db 585 LOKSASPDAAAGKSGLTPLHVAAYHYNQKVALLLDQOASPHAAAKNGYTPHIAAKNQ 644  
QY 712 DTODLL-----RGDAALLDAAKGCLARVKKLSPPD--VYNCRDQTQRHSTPLH 758  
Db 645 DIATLLLEYGADANAVTROGIVASVHLAAQEGHVDVMSLLGRNANVNLNKS--LTP 702  
QY 759 LAAGYNNLEVAEYLLQHGADVNAQDKGLIPLHNAASGYCHVDVAALLIKYNAVNATDKW 818  
Db 703 LAAQEDRVNVAEVLVNOGAHVDAQTKMGYTPPHVCHGYGNIKIVNFFLQHSKAVNAKT 762  
QY 819 AFTPLHEAAQKRTOLCALLAHGADPTLKNOEGQTPLDL-----VSADDVSALIT-AA 871  
Db 763 GYTPLHQAQOQTHIINVLLQNNASPNELTVNGNTALGTRRLGYISVVDTLKIVTEET 822  
QY 872 MPSPALPCYK---POVLNGVRSFGATADALS GSPSPSSLSAASSLDNLSGFSSELSSV 928  
Db 823 MTTTIVTEKHKMNPETMNEVLD-----MSDDEVKANAPEMLSDG----- 863  
QY 929 VSSSGTEGASSLEKKEKVPVGVDFSIIOFVRNLGLEHLMDFEREQ-ITLDVLMGHKELK 987  
Db 864 -----EYISDVEEGEDAMTGTDKYLGPDQDK 890  
QY 988 EIG 990  
1:1

```

Db      891 ELG 893

RESULT 6
Q61307
ID      Q61307      PRELIMINARY;      PRT: 1943 AA.
AC      Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; Q08866; Q08867;
DT      01-JUN-1998 (TrEMBLrel. 06, Created)
DT      01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT      01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE      Ankyrin-3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin-3).
OS      ANK3.
GN      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RS      SEQUENCE FROM N.A. (ISOFORMS 1-6).
RP      STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX      MEDLINE=95340633; PubMed=7615634;
RA      Peters L.L., John K.M., Lu F.M., Eichler E.M., Higgins A., Yialamas M.,
RA      Turtzo L.C., Otsuka A.J., Lux S.E.;
RT      "Ank3 (epithelial ankyrin), a widely distributed new member of the
RT      ankyrin gene family and the major ankyrin in kidney, is expressed in
RT      alternatively spliced forms, including forms that lack the repeat
RT      domain.";
RL      J. Cell Biol. 130:313-330(1995).
RN      [2]
RS      SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RP      STRAIN=C3H/HEJ; TISSUE=BONE MARROW;
RX      MEDLINE=97213781; PubMed=9060470;
RA      Hooch T.C., Peters L.L., Lux S.E.;
RT      "Isoforms of ankyrin-3 that lack the NH2-terminal repeats associate
RT      with mouse macrophage lysosomes.";
RL      J. Cell Biol. 136:1059-1070(1997).
CC      -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
CC      SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
CC      BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
CC      AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
CC      DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
CC      THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC      THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC      -!- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC      ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC      -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1 (SHOWN HERE), 2, 3,
CC      4, 5, 6, 7 AND 8; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC      -!- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC      EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC      LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC      SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC      -!- SIMILARITY: CONTAINS 24 ANK REPEATS.
DR      EMBL; L40631; AAB01603.1; -
DR      EMBL; L40632; AAB01605.1; -
DR      EMBL; U92275; AAB58381.1; -
DR      EMBL; L40633; AAB01602.1; -
DR      EMBL; L40632; AAB01604.1; -
DR      EMBL; L40632; AAB01606.1; -
DR      EMBL; L40632; AAB01607.1; -
DR      EMBL; U92274; AAB58380.1; -
DR      HSSP; P55273; 1B18.
DR      MGD; MGI:88026; Ank3.
DR      InterPro; IPR002110; ANK.
DR      InterPro; IPR000488; Death.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00023; ank; 24.
DR      Pfam; PF00531; death; 1.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01415; ANKYRIN.
DR      SMART; SM00248; ANK; 21.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50088; ANK_REPEAT; 21.
DR      PROSITE; PS50297; ANK_REPEAT_REGION; 1.

DR      PROSITE; PS50017; DEATH_DOMAIN; 1.
KW      Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW      Multigene family.
FT      DOMAIN 1 826 ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT      DOMAIN 23 810 24 X ANK MOTIF REPEATS.
FT      DOMAIN 856 1437 SPECTRIN BINDING DOMAIN.
FT      DOMAIN 1438 1943 REGULATORY DOMAIN
FT      (REGULATES BINDING OF ANKYRIN TO SPECTRIN
FT      AND THE BAND 3 PROTEIN).
FT      REPEAT 23 55 ANK MOTIF 1.
FT      REPEAT 56 88 ANK MOTIF 2.
FT      REPEAT 89 121 ANK MOTIF 3.
FT      REPEAT 122 154 ANK MOTIF 4.
FT      REPEAT 155 183 ANK MOTIF 5.
FT      REPEAT 184 216 ANK MOTIF 6.
FT      REPEAT 217 249 ANK MOTIF 7.
FT      REPEAT 250 282 ANK MOTIF 8.
FT      REPEAT 283 315 ANK MOTIF 9.
FT      REPEAT 316 348 ANK MOTIF 10.
FT      REPEAT 349 381 ANK MOTIF 11.
FT      REPEAT 382 414 ANK MOTIF 12.
FT      REPEAT 415 447 ANK MOTIF 13.
FT      REPEAT 448 480 ANK MOTIF 14.
FT      REPEAT 481 513 ANK MOTIF 15.
FT      REPEAT 514 546 ANK MOTIF 16.
FT      REPEAT 547 579 ANK MOTIF 17.
FT      REPEAT 580 612 ANK MOTIF 18.
FT      REPEAT 613 645 ANK MOTIF 19.
FT      REPEAT 646 678 ANK MOTIF 20.
FT      REPEAT 679 711 ANK MOTIF 21.
FT      REPEAT 712 744 ANK MOTIF 22.
FT      REPEAT 745 777 ANK MOTIF 23.
FT      REPEAT 778 810 ANK MOTIF 24.
FT      VARSPPLIC 1 805 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT      VARSPPLIC 1 849 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT      VARSPPLIC 833 853 MISSING (IN ISOFORM 7 AND ISOFORM 8).
FT      VARSPPLIC 834 855 VKRAPEKLSDGEIISDGEEG -> GDCKTWEKIPKQVEV
FT      VARSPPLIC 850 855 SDGEEG -> MALPHS (IN ISOFORM 2 AND
FT      VARSPPLIC 852 852 G -> GDCKTWEKIPKQVEVLYKS (IN ISOFORM 4
FT      VARSPPLIC 1569 1764 AND ISOFORM 6).
FT      VARSPPLIC 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;
SQ      SEQUENCE 1943 AA; 212013 MW; B23C4B6F82904804 CRC64;

Query Match 13.8%; Score 891; DB 11; Length 1943;
Best Local Similarity 28.0%; Pred. No. 1.3e-46;
Matches 339; Conservative 167; Mismatches 495; Indels 209; Gaps 39;

QY 68 AVAAARIMSGRRCCAGGACASAAAEVPAARELFECACRNGDVERVKRLVTPKVNRSR 127
DB 10 AKPAHRKRKGGKSDANASYLRAARAGHLEKA----LDYIKNG-----VDVNICN 54
QY 128 TAGRKSTPLHFAAGFGKQDVVEYLLONGANVQARDGGLIPLHNACSFGEAEVNVNLLRH 187
DB 55 QNGLNA--LHLASKEGHEVVEVSELLQREANVDAATKGNLTALIASLAGQAEVVKVLVTN 112
QY 188 GADPNARDNNVYTPHLEAAIKGKIDVCIIVLHOGAETPIRNTDGTALDLA-----DPSA 242
DB 113 GANVNAQSONGFTPLYNAAQENHLEVVVRLDNGASOSLATEGFTPLVALQOQHQQV 172
QY 243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
DB 173 SLLLENDTKGVRLPALHIAARKDDTKAAALLQNDNTNADVESKSGFTPLHIAAHYGNIN 232
QY 300 IVQVLLQHGADYHAKDKGLVPLHNACSYGHEYTELLVKHGACVNMADLWQFTPLHEAA 359
DB 233 VATLLNRAAADVFTARNDDITPLHVASKRGNANNVKILLDRGAKIDAKTRDGLTPLHCGA 292
QY 360 SKNREVCSSLISYGAADPTLLNCHNKSAIDLAPTOLKERLAYEFKGHSLIQAAREADVT 419
DB 293 RSGHEOVVEMLLDRSA-PILSKTKN-----GLSPLHMATQGD-- 328

```



QY 655 HGAVVNVADLWKFTPLHAAAGKYEICKLLOHGADPTKKNRDGNTPDL-VKDGDTDI 713  
 Db 581 KSASPDAGKSGTLPLHVAHYDNQKVALLLDQASPHAAAKNGYPLHHTAAKNQMDI 640  
 QY 714 QDL-----RGDAALDAAKGGCLARYKYLSSPD-NVNCRTDQGRHSTPLHLA 760  
 Db 641 ATSLLEYGADANAVTRQGIASVHLAAQEGHVDVMSLLSRANVNLNKRGLN-PLHLG 698  
 QY 761 AGYNNLEVAEYLLQHGADVNAQDKGGLPLHNAASYGHVDVAAALLIKYNACVNATDKWAF 820  
 Db 699 QGEDRVNVAEYLVNQGAHVDAQTKMGYTPHVGCHYGNIKIVNLFQHSKYNATKNGY 758  
 QY 821 TPLHAAQKQRTQCALLLAHGADPTLKNQEGQTPDLVS-----ADDVSALITAAAMP 873  
 Db 759 TPLHQAQQGHTHIIIVLLQNNASPNELTVNGNTALALARRLGYISVVDTLKVVTEETM 818  
 QY 874 PSALPSCYK---POVLNGVRSFGATADALSSGSPSSLSAASSLDNLSGFSSELSSVVS 930  
 Db 819 TTTITERKHKMNPETMNEVLDD--MSDDEVGKG-SAPEKLSHGEXI----- 860  
 QY 931 SSGTEGASSLEKKEVPGVDFSIOTFVRNLGLEHLMDFEREQITLDVLVEMGHKELKEIG 990  
 Db 861 SDGEEG-----EDAITGDTOKSLGPQDLKEIG 887

## RESULT 8

O70511 PRELIMINARY; PRT; 2622 AA.  
 ID O70511; AC O70511; RA 01-AUG-1998 (Tremblrel. 07, Created)  
 DT 01-AUG-1999 (Tremblrel. 11, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE 270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).  
 GN ANK3  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98417633; PubMed=9744885;  
 RA Zhang X., Bennett V.;  
 RT "Restriction of 480/270-kD ankyrin G to axon proximal segments  
 requires multiple ankyrin G-specific domains."  
 RL J. Cell Biol. 142:1571-1581(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Carpenter S.S., Zhang X.;  
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 934-1220 FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SKELETAL MUSCLE;  
 RX MEDLINE=98330444; PubMed=9664041;  
 RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;  
 RT "Ankyrin G is associated with the postsynaptic membrane and the  
 sarcoplasmic reticulum in the skeletal muscle fibre."  
 RL J. Cell Sci. 111:2197-2207(1998).  
 DR EMBL; AF102552; AAC78143.1; -;  
 DR EMBL; AF065150; AAC18853.1; -;  
 DR HSSP; P55273; IBI8.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00023; ank; 24.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00791; ZU5; 1.  
 DR SMART; SM00248; ANK; 21.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00218; ZU5; 1.  
 DR PROSITE; PS50088; ANK\_REPEAT; 21.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 KW ANK repeat; Repeat.

FT NON\_TER 2622 2622  
 SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;  
 Query Match 13.7%; Score 885.5; DB 11; Length 2622;  
 Best Local Similarity 30.5%; Pred. No. 4.6e-46;  
 Matches 283; Conservative 124; Mismatches 373; Indels 147; Gaps 24;  
 QY 105 ACRNGDVERVKRLVTPKVNRSRTAGRK-STPLHFAAGFRGKDVVYLLQNGANVQARDD 163  
 Db 81 ASKEGHEVVSSELLQREA--NVDAATKKGNTALHIALAGAEVVKVLYVINGANVNAQSQ 138  
 QY 164 GGLIPLHNACSGHAEVNNLLLRHGADPNARDNNWYTPLHEAAIKGIDVICVILLQHGAE 223  
 Db 139 NGFTPLYMAAQENHLEVRFLDNGASQSLATEDGETPLAVALQQGHQDVVSLLENDTK 198  
 QY 224 PIRNTDGTALDAPSAKAVLTGEYKDELLESARSNGNEEKMALLTPLNVNCHASDG 283  
 Db 199 GKVRUPALHIAAKDDTKAAALL-----LQNDTNDIESKVV-----VNRATESG 243  
 QY 284 RKSTPLHLAAGYNRVKIVQLLQHGADVHAKDKGLVPLHNACSYGHEVTELVKKGAC 343  
 Db 244 --FTSLHIAAHYGNINVTALLNRAAADFTARNITPLHVASKRGANVMVKLLDRGAK 301  
 QY 344 VNAMDLMQFTPLHEAASKNRVVCSSLLSYGADPTLNCNKSALIDAPTQPKERLAYE 403  
 Db 302 IDAKTRDGLTPLHCGARSGHEQVEMLLDRAA-PILSKTKN----- 341  
 QY 404 FKGHSLLOAREADVTRIKKHSLEVMNFKH-----PQTHETALHCAASPYPKRKQI 456  
 Db 342 --GLSPLHMATQGD-----HLMCVQLLIQHNVPVDDVTNDYLTALHVAHCHGYK---V 390  
 QY 457 CELLRKGANINEKTRFETPLHVASEKAHNDVVVVVKEAKVNALDNLGQTSLSHRAAY 516  
 Db 391 AKVLLDKKANPAKALNGFTPLHIAKKNRIRVWELLKHGASIQAVTESGLTPIHVAAF 450  
 QY 517 CQHLQTCRLLSYGCDPNIIISLOGFTALQM-----NENVOOLLQEGISLNSADRO-- 569  
 Db 451 MGHVNTIVSOLMHGASPNNTNVRGETALHMAARSGQAEVRYLVODGAQVEAKAKDDQTP 510  
 QY 570 LLEAAKAGDVTYVKLCITV--QSVNCRDIEGRQSTPLHFAAGYNRVSVVVEYLLQHGADVHA 628  
 Db 511 LHISARLGRADIVQOLLOOGASPNAAATSG--YTPLHSAREGHEDVAFAFLDHGASLSI 568  
 QY 629 KDKGGLVPLHNACSYGHEVAELLVKGAVNVNADLWKFTPLHAAAKGKYEICKLLOH 688  
 Db 569 TTKKGFTPLHVAAYKGLVASULLQKASPDAAAGKSGLTPLHVAHYDNQKVALLLDQ 628  
 QY 689 GADPTKKNRDGNTPDL-VKDGDTDIQDL-----RGDAALLDAAKKGCLARY 735  
 Db 629 GASPHAAAKNGYTPHIAAKKNQMDIATSLLEYGADANVPVTRQGIASVHLAAQEGHVDV 688  
 QY 736 KKLSSPD-NVNCRTDQGRHSTPLHAAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 794  
 Db 689 SLLSRANVNLNKSQ--LTPHLAAQEDRVNVVAEVLVNOGAHVDAQTKMGYTPHLYGC 746  
 QY 795 SYGHVDVAAALLIKYNACVNATDKWAFPLHEAAQKRTQCALLLAHGADPTLKNQEGQ 854  
 Db 747 HYGNIKIVNLFQHSKYNATKNGYTPHQAQOQGHTHIIVNLLQNNASPNELTVNGNT 806  
 QY 855 PLDLVS-----ADDVSALLTAAMPSPALPSCYK---POVLNGVRSFGATADALSSG 904  
 Db 807 ALAIARRLGYISVVDTLKVVTEETITTTITERKHKMNPETMNEVLDD-----M 854  
 QY 905 PSSSLASAASLDNLS-GSFSSELSSVSSSGTEGASSLEKKEVPGVDFSIOTFVRNLGLEH 963  
 Db 855 SDDEVGKASAPEKLSDGEY-----ISDGEEG----- 880  
 QY 964 LMDIFEREQITLDVLVEMGHKELKEIG 990  
 Db 881 -----EDAITGDTOKSLGPQDLKEIG 901  
 RESULT 9





[illegible]

DR PROSITE; PS50088; ANK\_REPEAT; 21.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
KW ANK repeat; Repeat.  
SQ SEQUENCE 1549 AA; 170167 MW; 3ECEC852B148DFF CRC64;

Query Match 13.4%; Score 868.5; DB 5; Length 1549;  
Best Local Similarity 27.4%; Pred. No. 2.3e-45;  
Matches 311; Conservative 173; Mismatches 491; Indels 159; Gaps 34;

QY 98 AARELFACRNGDVERVKRLV---TPEKVSRSRTAGRKSTPLHFAAGFGRKDVVEYLLQN 154  
Db 38 ATTSFLRAARGDGLIKYVDFDLCGEISDINSCNANGLNA--LHLAAKDGYYDICECLLR 95

QY 155 GANYQARDGGLIPLHNACSFHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266  
Db 156 RTLLANGANPSLSTEDGFTPLAVAMQOCHDKIVAVLLENDVGRKVRPLPALHIAAKNDVN 215

QY 267 MMALLTPLNVNCHASDGRKSTPLHACGYHAEVNNLLRHGADPNARDNNWYTPLHEAAIKGKIDVC 214  
Db 96 GIKIDNATKGNATLHATSLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

QY 215 IVLLOHGAETIWTNRTDGRDLDLA----DPSKAVLTGEYK---KDELLSARSGNEEK 266



Db	216	AKLLQHPDNDADIVSKSGFTPLHIAAHYGNVDIATLLNNKADVNVYAKHNITPLHVAC	275
Qy	327	SYGHEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCISLLSYGADPTLLNCHN-K	385
Db	276	KWGLSICLTLLCRGAKIDAATDGLTPLHCASRSHVEVIRKHLQONA-PILTCTKNGL	334
Qy	386	SATDLAPTQPKERLAYEFKGHSLLOAAREADVTRIKKHSLE-----MNVF	432
Db	335	SALHMAAQGE-----HDEAHILLDNKAPVDEVYITLHVAACHGVKAKILLDDY	388
Qy	433	K-HPOTHE-----TALHCAASAPYKPKQICELLRLKGANINEKTKFELPLHVASEKARN	487
Db	389	KANPNARALNGFTPLHACKK---NRKKVELLIKGANIGATTESGLPLHVASFPGCI	445
Qy	488	DVEVVVYKHEAKVNALDNLGTSLSHRAAYCGHLOTCLRLLSYGCDDNITISLQGFAL---	544
Db	446	NIVYLLQHEASADLPTIRGETPLHAAARANQADIIRILL-SAKVDAIVREGQTPLHVA	504
Qy	545	QMGNEN-VQOOLQEGISLGNSEADR--OLLEAAKAGDVETVKKLCCTVQSVNCRDTEGRQ	600
Db	505	SRIGNINIIMLLQHGAEINAQSNKYSAHTAAKEQENIVOVLLLENGAEN-NATVTKG	563
Qy	601	STPLHFAAGYNRVSVYLLQHGADVHAKDKGGLVPLHNAACSYCHYEAEVALLVKHGAVN	660
Db	564	FTPLHLACKYKQNVQILLONGASIDFGKNDVTPLVHATHYNNPSIVELLKNGSSPN	623
Qy	661	VADLWKFPTPLHEAAAKYKIECKLLQHGADPTKKNRDNTPDL-VKGGDTDIQDLRLG	719
Db	624	LCARNGQCAITHIACKKNYLETAMQLQHGADVNIISKSGFSPHLAAQGNVDMVQLLE	683
Qy	720	DAALLDAKGCCLARVKLSSPDVNCRDQTGRHSTPLHAAAGVNNLEVAEYLLQHGADV	779
Db	684	YGVISAANKNL-----TPLHVAQEGHVLVSQILLEHCANI	720
Qy	780	NAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGTQLCALLL	839
Db	721	SERTNGVTPPLHMAHYGHLDLVKFFIENDADIEMSSNIGYTPPLHQAQOQGHIMIINLL	780
Qy	840	AHGADPTLKNOEGOTPLDVSADDSALLTAAMPSPALPCKYKQVPLNG-----V	889
Db	781	RKANPNALTQKGNATLHIAASN-----LGVTVWESLKIVTSTVINSNIGATEELKV	834
Qy	890	RSPGATADALSGSPSSLSAAS-----SLDNLSGSFSELSSVSSSGTE-----GAS	938
Db	835	MTPELWQFTLLSDSDDECDLLDHNYKYMATDCLKANYQDQKQNFDTTNDHLDVDS	894
Qy	939	SLEKKEVPGVDFSTQFVRNLGLHLM-D--IFEREQITLDVLVEMGHKELKEIGINAYGH	996
Db	895	VLNKKEILPNEMSCITELTE---IGHKPDNVVIARSOVHLGLVSF---LVDARGGSMRGY	948
Qy	997	RHKLKGE-----RLISGOGLNPVLTNTSGSGTILIDLSPDDKEFQS--	1041
Db	949	RINGVRIIVPKACAEPTITCRIVKPVORVNPPLMEGEALVSRILEMSPVDGMFLSPI	1008
Qy	1042	-VEEEMQSTVREHROGGHAGGIFNRYNILKIQVCNKKLWERYTHRRKEVSEENHNANE	1100
Db	1009	TLEVPHYGTLRK-----NEREII-ILRSDNGESWEH-NLYKDIIGEDINQTEE	1055
Qy	1101	RMLFGSPFNVAIIHKGFEDERHAYIGMFGAGIYFAENSSKNQVYVIG--GGT	1153
Db	1056	---FHSDRIVR-IVTQNPVPHFAV-----SRVQEVHVHPDGGT	1092
RESULT 14			
ID	P97582	PRELIMINARY:	PRT: 843 AA.
AC	P97582;		
DT	01-MAY-1997 (Tremblrel. 03, Created)		
DT	01-MAY-1997 (Tremblrel. 03, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	Ankyrin (Fragment).		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=WISTAR; TISSUE=BRAIN;
RA	Cosentino M.T., Jones O.T.;
RT	"Rat brain ankyrin membrane binding domain.;"
RL	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR	EMBL; U65916; AAB47551.1; -
DR	HSSP; Q00420; IAWC.
DR	InterPro: IPR002110; ANK.
DR	Pfam: PF00023; ank; 24.
DR	SMART; SM00248; ANK; 21.
DR	PROSITE; PS50088; ANK_REPEAT; 20.
DR	PROSITE; PS50297; ANK_REPEAT; 1.
KW	ANK repeat; Repeat.
FT	NON_TER 1
FT	NON_TER 843
FT	NON_TER 843
SQ	SEQUENCE 843 AA; 89982 MW; DC5A6AC78825D223 CRC64;
Query Match	
Best Local Similarity 13.3%; Score 858; DB 11; Length 843;	
Matches 275; Conservative 148; Mismatches 367; Indels 140; Gaps 23;	
Qy	105 ACRNGDVER-VKRLVTPPEKVNRSRTAGRKSTPLHFAAGFGRKDVVEYLLONGANVQARD 163
Db	3 AARAGNLKDVVEYLGKGIINTCQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 60
Qy	164 GGLIPLHNAACSGHAEVNVNLLRHGADPNARDNWNTPLEHAAIKGKIDVCIVLQHGAE 223
Db	61 KGNALHIAASLAGOAEVVKVLEKGANINAQSONGFTPLYMAAQENHIDVVKYLLENGAN 120
Qy	224 PTIRNTDGR-----ALDLADPSAKAVLTCEYKDEL---LESARSGNEEKEMALLTPLN 275
Db	121 QSTATEDGFTPLAVALQOQHQAVALLENDTKGVRLPALHIAARKDDTKSAALLQ-- 178
Qy	276 VNCHASDGRKS-----TPLHAAAGYNRVKIVQLLQHGADVHAKDKGDLVPLHN 324
Db	179 -NDHNAQVDSKMMVNRSTESGFTPLHIAAHYGNVNATLLNARGAAVDFTARGITPLHV 237
Qy	325 ACSYGHVEVTELLVKGACVAMDLWQFTPLHEAASKNRVEVCISLLSYGADPTLLNCHN 384
Db	238 ASKRGNTNMVKLLDRGQIDAKTRGLTPLHCAASRGHDQVVELLGERG-PLLARTKN 296
Qy	385 KSADLAPTQPKERLAYEFKGHSLLOAAREADVTRIKKHSLEVMNFKHP-----QTHE 439
Db	297 -----GLSPLHMAAQGDHVECVKHL---LQHKAPVDVDTLDYL 331
Qy	440 TALHCAASAPYKPKQICELLRLKGANINEKTKFELTPLHVASEKAHNDVVEVVKHEAK 499
Db	332 TALHVA--HCGHYRVTKLLDKRANPNARALNGFTPLHIAACKNRKIKWELLVKYIGAY 388
Qy	500 VNALDNLGOTSLHRAAYCGHLOTCLRLLSYGCDDNITISLQGFALQMGNEVQOILLQEG 559
Db	389 IQAITESGLTPIVPAFMGHLNIVLLLONGASPDVNTINIRGETALHM----- 435
Qy	560 SLGNSEADROLLEAAKAGDVETVKKLCCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
Db	436 -----AARAGEVEVVR--CLLRNGALVDARAREEQTPHLIASRLSKTEIVQL 480
Qy	619 LLQHGADVHAKDKGGLVPLHNAACSYCHYEAEVALLVKHGAVNVDLWKFTPLHEAAAKGK 678
Db	481 LLQHMHPDAATNGYTPPLHISAREGQVDVSVLLLEGAHAHSIATKGGFTPLHVAAYKGS 540
Qy	679 YEICKLLQHGADPTKKNRDNTPDLVLDKGGDTDIQDLRLGDAALLDAAKGCLARVKKL 738
Db	541 LDVAKLLLQRRAAADSAGKGLTPLHVAHYDN-----QKVALLLLEKG 584
Qy	739 SSPDNVNCRDQTGRHSTPLHIAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 798
Db	585 ASP-----HATAKNGYTPPLHIAAKKNQWQIASTLLNTGAETNTVTKQGVTPPLHLSQEGH 639
Qy	799 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGTQLCALLLHAGADPTLKNQEGQTPDL 858

Db 640 TDMVTLVLEKGANIHMSTRKSGTSLTSLHAAEDKVNVDILTKHGADODDAYTKLGYTPLLIV 699  
 QY 859 V-----SADDSVALLTAAMPSSA-LPSCYKP-----QVLNGVRSFGATADALSSGP 903  
 Db 700 ACHYGKVNKYNELLKOGANVNKATKNGYTPHLHQAQQGHTHTIINVLLQHCAPNATTA-- 757  
 QY 904 SSPSLSAASSLDNLSSGFSFSSSVSSSGTEGASLEKKEKPGVDFTSTQFVRNGLDEH 963  
 Db 758 NGNTALAIKRLGYI--SVVDTLKVVTEVTTTTTITTEKHLNAPETMTE----- 806  
 QY 964 LMDIFERE---OITLDVLVEMGHKELKEIG 990  
 Db 807 VLDVDEEGDDTVTGDGGEYLRPEDLKELG 836

## RESULT 15

Q17487 PRELIMINARY; PRT; 1809 AA.  
 AC Q17487;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE E. elegans ankyrin-related unc-44 (GB:U21734).  
 GN UNC-44.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinon-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gattung S.;  
 RT "The sequence of C. elegans cosmid B0350.";  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U50071; AAA93446.1; -.  
 DR HSSP; P42773; 1IHB.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR001360; GH\_1.  
 DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00023; ank; 24.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00791; ZU5; 1.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR SMART; SM00248; ANK; 21.  
 DR SMART; SM00005; DEATH; 1.  
 DR PROSITE; PS50088; ANK REPEAT; 22.  
 DR PROSITE; PS50297; ANK REP REGION; 1.  
 DR PROSITE; PS50017; DEATH DOMAIN; 1.  
 DR PROSITE; PS00572; GLYCOSTYL\_HYDROL\_F1\_1; UNKNOWN\_2.  
 KW ANK repeat; Repeat.  
 SQ SEQUENCE 1809 AA; 198193 MW; AD86349AA64534F8 CRC64;

Query Match 13.3%; Score 858; DB 5; Length 1809;  
 Best Local Similarity 26.3%; Pred. No. 1.3e-44;  
 Matches 319; Conservative 163; Mismatches 443; Indels 286; Gaps 39;  
 QY 91 AAFAVEPAARE---LFEACRNGDVERVRVLY--TPKVNSTRDTAGRKSTPLHFAAGFGRK 145  
 Db 22 AAPAEPGPAEGSASFLRAARAGDLEKLELRAGTDINTSNAGLNS--LHLASKEGHS 79  
 QY 146 DYVEYLLQGANVQARDGGGLPLHNACSGHAEVNNLLRHGADPNARNNNYTPHHEA 205  
 Db 80 EVVRELIRKQAOVDAATRGNTALTALHIASLAGOSLIVILVENGANNVOSVNGFTPLYMA 139  
 QY 206 ATKKIDVICVILQHGAEPTINTDGTALDLADPSAKAVLTGEYKDELLESARSNEE 265  
 Db 140 AOENHEEVVYKLLKHGANQALSTEDGFTPLAVA-----LQGHDR 179  
 QY 266 KMWALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLQHGADVHAKDKGLVPLHN 324  
 Db 180 VYAVLLE-----NDSKGVRLPALHIAAKKDDTTAATLLQNEHNPVDTSKSGFTPLHI 233  
 QY 325 ACSYGHYEVTLLVKGACVNMADLWQFTLHEAAKKNRVEVCSLLLSYGADPTLLNCHN 384  
 Db 234 AAHYGHENVGQLLEKGANVYQARHNISPLHVATKWTNNANLLLSRGA---IIDSRT 290  
 QY 385 KSAIDLAPTPOLKERLAYEFKCHSLLOQAAREADVTRIKKHLISLEVMNFKHPOTHTALHC 444  
 Db 291 KDLL-----TPLHC 299  
 QY 445 AAAPYKPKKOICELLRRKANINEKTEFTPLHVASEKANDVVEVVVHKAQVN--A 502  
 Db 300 AARSQH---DOVVDLLVVQGAIPISAKTKNGLAPLHMAAGDHVDAARTLLYHRAPVDDVT 356  
 QY 503 LDNLGQTSLHRAAYCGHLOTCLLSYGCDPNIISLOGFTALQMGNE----- 550  
 Db 357 VDYL--TPLHVAACHGVRAKLLDRSADPNRSLNGFTPLHITACKKRIKVVLELLKY 414  
 QY 551 ---VQQLLOEGIS-----LGNSEADRLLEAKAGDVEIVKCLTVOVNCRDIEGRQ 600  
 Db 415 RAAEATESTESGLTPLHVAAFMGAINIVYLQOAGNPVDET-----RG 458  
 QY 601 STPLHFAAGYNRVSVVEYLQHGADVHAKDKGLVPLHNACSYGHEVAELLVKGAVVN 660  
 Db 459 ETPLHFAARANTDVRVLRNGAKVDAQARELOTPLHIAKRLGNTDIVILLQGANSN 518  
 QY 661 VADLWKEFTPLHFAAKGKEICKLLLOHGADPTKKNRDGNTPLDLV-KGDTDTIDQLL-- 717  
 Db 519 ATTRDNYSPHLHAAKEGEEVAGILLDHNADKTLTKKGFPLHASKYGNLEVVRLLE 578  
 QY 718 RGDAAALLDAARKGCLARVKYKLLSSPDNVNCRDTCGRHS--TPLHLAGYNNLEVAEYLLQHG 776  
 Db 579 RGTVP-----DIEGKNQVTPHVAHNNNDKVMALLJENG 613  
 QY 777 ADVNNAQKGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFPTPLHFAAKGRTQLCA 836  
 Db 614 ASAKAAAKNGYTPHLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673  
 QY 837 LLLAHGADPTLKNOEGOTPLDLYSADD-----VSALLTAAMPSPSALPSCYK 882  
 Db 674 LLTENGSDVGAKANNGLTAMHLCAQEDHPVPAQILYNGAEINSKTNAGYTPLHV-ACHF 732  
 QY 883 POVLNGVR---SPCATADALSSGSPSSLSAASSLSGTSFSELSSVSVSSSGTEGASS 939  
 Db 733 GQ-LNMVKFLVENGADVGEKTRASYTPHLHQAQOQHNNVCVRYLLE-----NCASP 781  
 QY 940 LEKKEVPGVDFSTQFVRNLG-----LEHLMDFEREQITLDVLVEMGHKELKEIGIN-- 992  
 Db 782 NEQTATGQTPLSIAQ---RLGYVSVVETLTVTETVTITETTVTDERYKQPNEMMETM 838  
 QY 993 -----AYGHRHKL-----KGVRLISGOOGLNPYLTNTSG--- 1024  
 Db 839 FSESEDEGQAHEAHAHEKDFSDNLQTGLQDSTGVHMTHTGEQLLQSRQSELENGGAIP 898

Wed Feb 12 16:08:54 2003

QY 1025 ----SCTILIDLSPDDKEFQSVEEEMQS-----TVREHRDG--CHAGGTFRNYNLIKOKVC 1075  
Db 899 KINSQ---GMSQ-EKEFAKIAPVATSSPIATSSNSQSFQIAPRAGSISGQFQQOQLHGAG 953  
QY 1076 NKKLWERYTHRKE--VSEENHNHANERMLPHGS-----PFVNAIIHKG-----DERHA 1123  
Db 954 PEDNLEELVRRQNHPIINAGNYDNGGVAMLENGHADNVPIGHVHTQPSFLISFLVDARGG 1013  
QY 1124 YIGGMFGAGIYFAENSXSNQYVYIGGGTGCPCVHKDRSCYICHRQLLFCRVTLGKSFQ 1183  
Db 1014 AMRGCRHSGVRIIVPPRKSQ-----PIR-----VTCRY-LRKD--- 1046  
QY 1184 FSAMKMAHSPP 1194  
Db 1047 ----KLAHPPP 1053

Search completed: February 12, 2003, 07:44:28  
Job time : 50.6325 secs

